SLR:kdn 03/25/02 2847-62447 106961 doc

EXPRESS MAIL LABEL NO. EV059538951US
DATE OF DEPOSIT: 3/25/02

JC05 Rec'd PCT/PTO 25 MAR 2002

FORM PTO-1390 U S DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY'S DOCKET NUMBER
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. § 371	2847-62447
	U S APPLICATION NO (If known, see 37 C F R § 1 5) Not yet assigned 10/089211
INTERNATIONAL APPLICATION NO. PCT/US00/27210	INTERNATIONAL FILING DATE October 2, 2000
	PRIORITY DATE CLAIMED October 1, 1999

TITLE OF INVENTION
MANNOSIDASES AND METHODS FOR USING SAME

APPLICANT(S) FOR DO/EO/US
William E. Hintz and Caleb Joshua Eades

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1 ☒ This is a **FIRST** submission of items concerning a filing under 35 U S.C § 371.

2 ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U S C § 371.

3. ☒ This is an express request to begin national examination procedures (35 U.S C. § 371(f) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U S.C. § 371(b) and PCT Articles 22 and 39(1).

4 ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.

5. ☒ A copy of the International Application as filed (35 U S C § 371(c)(2))

a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).

b. ☐ has been transmitted by the International Bureau

c. ☒ is not required, as the application was filed in the United States Receiving Office (RO/US).

6 ☐ A translation of the International Application into English (35 U.S.C. § 371(c)(2)).

7 ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S C § 371(c)(3))

a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).

b. ☐ have been transmitted by the International Bureau.

c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.

d. ☒ have not been made and will not be made

8 ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S C. § 371(c)(3)).

9 ☒ An oath or declaration of the inventor(s) (35 U.S.C § 371(c)(4)).

10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. § 371(c)(5))

Items 11. to 16. below concern document(s) or information included:

11. ☒ An Information Disclosure Statement under 37 C.F R §§ 1.97 and 1.98.

☒ Copies of References Cited.

12. ☒ An assignment document for recording. A separate cover sheet in compliance with 37 C.F.R. §§ 3.28 and 3.31 and the Recordal fee of $40.00 is included.

13 ☒ A **FIRST** preliminary amendment.

☐ A **SECOND** or **SUBSEQUENT** preliminary amendment

14. ☐ A substitute specification.

15 ☐ A change of power of attorney and/or address letter.

16 ☒ Other items or information:

☒ Abstract on a separate page.
☒ Written Opinion.
☒ International Search Report.
☒ Sequence Listing (Paper copy, disk, Statement in Compliance)

24197

[redacted]

SLR:kdn 03/25/02 2847-62447 106961.doc

EXPRESS MAIL LABEL NO. EV059538951US
DATE OF DEPOSIT: 3/25/02

U.S. APPLICATION NO (if known, see 37 C F R § 1 5)	INTERNATIONAL APPLICATION NO	ATTORNEY'S DOCKET NUMBER
Not yet assigned 10/089211	PCT/US00/27210	2847-62447

	CALCULATIONS (PTO USE ONLY)	
17. ☒ The following fees are submitted		
BASIC NATIONAL FEE (37 C.F.R. §§ 1.492(a)(1)-(5)):		
Neither International Preliminary Examination fee (37 C.F R § 1.482) nor International Search fee (37 C.F R. § 1 445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO **$1,040.00**		
International Preliminary Examination fee (37 C F R. § 1 482) not paid to USPTO but International Search Report prepared by the EPO or JPO **$890.00**		
International Preliminary Examination fee (37 C F.R. § 1.482) not paid to USPTO but International Search fee (37 C.F.R. § 1.445(a)(2)) paid to USPTO as an International Searching Authority **$740.00**		
International Preliminary Examination fee paid to USPTO (37 C.F R § 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4) **$710.00**		
International Preliminary Examination fee paid to USPTO (37 C F R. § 1 482) and all claims satisfied provisions of PCT Article 33(1)-(4) **$100.00**		
ENTER APPROPRIATE BASIC FEE AMOUNT =	$	**$710.00**
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 C F R. § 1.492(e)).	$	

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	19 – 20 =	0	x **$18.00**	$	**0.00**
Independent Claims	2 – 3 =	0	x **$84.00**	$	**0.00**
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ **$280.00**	$	

TOTAL OF ABOVE CALCULATIONS =	$	**0.00**
☒ Reduction of 1/2 for filing by small entity. Small entity status is claimed for this application	$	**$355.00**
SUBTOTAL =	$	**0.00**
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 Months from the earliest claimed priority date (37 C F R. §§ 1 492(f)). +	$	
TOTAL NATIONAL FEE =	$	**0.00**
Fee for recording the enclosed assignment (37 C.F.R. § 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 C F R §§ 3.28, 3.31). **$40.00** per property. +	$	**$40.00**
TOTAL FEES ENCLOSED =	$	**$395.00**
REFUND →	$	
CHARGE →	$	

a. ☒ A check in the amount of $ 395 00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. _____ in the amount of $ _____ to cover the above fees. A duplicate copy of this sheet is enclosed

c. ☒ The Director is hereby authorized to charge any additional fees that may be required, or credit any overpayment, to Deposit Account No. 02-4550. A duplicate copy of this sheet is enclosed.

d. ☒ Please return the enclosed postcard to confirm that the items listed above have been received

NOTE: Where an appropriate time limit under 37 C.F.R. § 1.494 or § 1.495 has not been met, a petition to revive (37 C.F.R. § 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

KLARQUIST SPARKMAN, LLP
One World Trade Center, Suite 1600
121 S W Salmon Street
Portland, OR 97204-2988

SIGNATURE
Sheree Lynn Rybak, Ph D.
NAME

47,913
REGISTRATION NUMBER

cc: Docketing

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: William E. Hintz and Caleb Joshua Eades
Art Unit: Not yet assigned
Application No. Not yet assigned

Filed: herewith

For: MANNOSIDASES AND METHODS FOR USING SAME

Examiner: Not yet assigned

Date: March 25, 2002

COMMISSIONER FOR PATENTS
WASHINGTON, D.C. 20231

PRELIMINARY AMENDMENT

Please enter this amendment before the fees are calculated.

Please amend the specification as follows:

On page 1, line 2 under the title, insert the following paragraph:

--CROSS-REFERENCE TO RELATED APPLICATIONS

This is the National Stage of International Application No. PCT/US00/27210, filed October 2, 2000, which in turn claims the benefit of U.S. Provisional Application No: 60/157,341, filed October 1, 1999.--

Please amend the claims as follows. Claims that have not been amended are denoted as reiterated for the examiner's convenience.

1. (Reiterated) A purified protein having mannosidase activity, comprising an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence selected from the group consisting of SEQ ID NOS: 3, 6 and 18;

(b) an amino acid sequence that differs from those specified in (a) by one or more conservative amino acid substitutions; and

(c) amino acid sequences having at least 60% sequence identity to the sequences specified in (a).

2. (Reiterated) A specific binding agent, that binds to the purified protein of claim 1.

3. (Amended) An isolated nucleic acid molecule encoding the protein of claim 1.

4. (Amended) A recombinant nucleic acid molecule, comprising a promoter sequence operably linked to the nucleic acid of claim 3.

5. (Amended) A cell, transformed with the recombinant nucleic acid molecule of claim 4.

6. (Amended) The transformed cell of claim 5, wherein the cell is an insect cell, a yeast cell, an algae cell, a bacterial cell, a mammalian cell, or a plant cell.

7. (Amended) A transgenic fungus, comprising the recombinant nucleic acid of claim 4.

8. (Amended) A method for altering a glycosylation pattern of a macromolecule, comprising contacting the macromolecule with the purified protein of claim 1.

9. (Reiterated) The method of claim 8, wherein contacting the macromolecule is performed *in vitro*.

10. (Amended) A method for producing a macromolecule having an altered glycosylation pattern, comprising allowing the transformed cell of claim 4 to produce the macromolecule.

11. (Amended) An isolated nucleic acid molecule, comprising a sequence selected from the group consisting of:

(a) at least 20 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;

(b) at least 30 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;

(c) at least 40 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;

(d) at least 15 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;

(e) at least 20 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;

(f) at least 30 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;

(g) at least 40 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;

(h) at least 50 contiguous nucleotides of the sequence shown in SEQ ID NO: 4; and

(i) at least 50 contiguous nucleotides of the sequence shown in SEQ ID NO: 1.

12. (Amended) A method for altering a glycosylation pattern of a macromolecule in a sample, comprising:

(a) adding the purified protein of claim 1 to the sample;

(b) incubating the sample with the purified protein; and

(c) allowing the purified protein to hydrolytically remove at least one terminal mannoside residue from a macromolecule in the sample.

13. (Amended) A method for identifying a nucleic acid sequence encoding a mannosidase, comprising:

(a) hybridizing the nucleic acid sequence under high-stringency conditions to the nucleic acid molecule of claim 11; and

(b) identifying the nucleic acid sequence as one that encodes a mannosidase.

14. (Reiterated) A mannosidase identified by the method of claim 13.

Please add the following new claims:

15. (New) The purified protein of claim 1, wherein the amino acid sequence comprises at least 70% sequence identity.

16. (New) The purified protein of claim 1, wherein the amino acid sequence comprises at least 80% sequence identity.

17. (New) The purified protein of claim 1, wherein the amino acid sequence comprises at least 90% sequence identity.

18. (New) The purified protein of claim 1, wherein the amino acid sequence comprises at least 95% sequence identity.

19. (New) The method of claim 8, wherein contacting the macromolecule is performed *in vivo*.

REMARKS

By this amendment, claims 3-8 and 10-13 are amended and claims 15-19 are added. Therefore, claims 1-19 are now pending.

In addition, the cross-reference to related applications was added to the specification by this amendment.

Claims 3-5, 7, 8, 12 were amended to clarify the antecedent basis.

Claim 6 was amended to correctly depend from claim 5.

Claims 4, 6 and 10 were amended to remove unnecessary claim language.

Claim 11 was amended to clarify the claim. Support can be found on page 4, lines 19-23 and page 11, lines 1-3.

Claim 13 was amended to depend from claim 11 and to clarify the claim. Support can be found on page 11, line 29- page 12, line 28.

Claims 15-19 were added. Support for these claims can be found in the specification as follows:

Claims 15-18: page 7, lines 14-23; and

Claim 19: page 27, line 18- page 28, line 6.

Therefore, no new matter is added by this amendment.

If there are any questions regarding this amendment, please telephone the undersigned at the telephone number below.

Respectfully submitted,
KLARQUIST SPARKMAN, LLP

By _____
Sheree Lynn Rybak, Ph.D.
Registration No. 47,913

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391
Facsimile: (503) 228-9446

Marked-up Version of Amended Claims and Specification
Pursuant to 37 C.F.R. §§ 1.121(b)-(c)

In the specification, page 1, line 2 under the title, insert the following paragraph:

CROSS-REFERENCE TO RELATED APPLICATIONS

This is the National Stage of International Application No. PCT/US00/27210, filed October 2, 2000, which in turn claims the benefit of U.S. Provisional Application No: 60/157,341, filed October 1, 1999.--

In the claims:

3. (Amended) An isolated nucleic acid[,] molecule encoding [a] the protein [according to] of claim 1.

4. (Amended) A recombinant nucleic acid molecule, comprising a promoter sequence operably linked to [a] the nucleic acid [sequence according to] of claim 3.

5. (Amended) A cell, transformed with [a] the recombinant nucleic acid molecule [according to] of claim 4.

6. (Amended) The transformed cell of claim [6] 5, wherein the cell is [selected from the group consisting of:] an insect cell, a yeast cell, an algae cell, a bacterial cell, a mammalian cell, [and] or a plant cell.

7. (Amended) A transgenic fungus, comprising [a] the recombinant nucleic acid [according to] of claim 4.

8. (Amended) A method for altering [the] a glycosylation pattern of a macromolecule, comprising contacting the macromolecule with [a] the purified protein [according to] of claim 1.

10. (Amended) A method for producing a macromolecule having an altered glycosylation pattern, comprising[:

(a) providing a transformed cell according to claim 4; and

(b)] allowing the transformed cell of claim 4 to produce the macromolecule.

11. (Amended) An isolated nucleic acid molecule, comprising a sequence selected from the group consisting of:

(a) [at least 15 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;

(b)] at least 20 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;

[(c)] (b) at least 30 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;

(c) at least 40 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;

(d) at least 15 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;

(e) at least 20 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;

(f) at least 30 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;

(g) at least [15] 40 contiguous nucleotides of the sequence shown in SEQ ID NO: [17] 4;

(h) at least [20] 50 contiguous nucleotides of the sequence shown in SEQ ID NO: [17] 4; and

(i) at least [30] 50 contiguous nucleotides of the sequence shown in SEQ ID NO: [17] 1.

12. (Amended) A method for altering [the] a glycosylation pattern of a macromolecule in a sample, comprising:

(a) adding [a] the purified protein [according to] of claim 1 to the sample;

(b) incubating the sample with the purified protein; and

(c) allowing the purified protein to hydrolytically remove at least one terminal mannoside residue from a macromolecule in the sample.

13. (Amended) A method for [isolating] identifying a nucleic acid sequence encoding a mannosidase, comprising:

(a) hybridizing the nucleic acid sequence under high-stringency conditions to the nucleic acid molecule of claim 11 [at least 50 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 1, 4, and 17]; and

(b) identifying the nucleic acid sequence as one that encodes a mannosidase.

15. (New) The purified protein of claim 1, wherein the amino acid sequence comprises at least 70% sequence identity.

16. (New) The purified protein of claim 1, wherein the amino acid sequence comprises at least 80% sequence identity.

17. (New) The purified protein of claim 1, wherein the amino acid sequence comprises at least 90% sequence identity.

18. (New) The purified protein of claim 1, wherein the amino acid sequence comprises at least 95% sequence identity.

19. (New) The method of claim 8, wherein contacting the macromolecule is performed *in vivo*.

10/089211
10 Rec'd PCT/PTO 25 MAR 2002

MANNOSIDASES AND METHODS FOR USING SAME

Field of the Invention

The invention relates to mannosidases and methods of using mannosidases for altering the glycosylation pattern of macromolecules.

Background

The N-linked protein glycosylation pathways are fairly well characterized in higher eukaryotes (Kornfeld and Kornfeld, *Ann. Rev. Biochem.* **54**:631-664, 1985), however, less is known about such pathways in lower eukaryotes. What is known about protein glycosylation in lower eukaryotes has come largely from studies of the yeast *Saccharomyces cerevisiae* and may not be sufficient to describe N-linked glycosylation in other lower eukaryotes, such as filamentous fungi.

Evolutionary studies suggest that the filamentous ascomycetes diverged from yeasts from 400 million years ago (Berbee and Taylor, *Mol. Biol. Evol.* **9**:278-284, 1992) to 1 billion years ago, the latter being about the time the fungal branch split from plants and animals (Doolittle et al., *Science* **271**:470-477, 1996). Filamentous fungi produce N-glycan structures that appear to be different than those produced in yeast, suggesting a different mode of formation of these structures than in yeast. Studies have shown the presence in filamentous fungi of N-glycan structures containing five mannose units ($Man_5GlcNAc_2$), suggesting processing of a $Man_5GlcNAc_2$ precursor (Mares et al., *Eur. J. Biochem.* **245**:617-625, 1997; Chiba et al., *Curr. Microbiol.* **27**:281-288, 1993). The mechanisms of synthesis of high mannose N-glycans in filamentous fungi seem to differ from corresponding mechanisms in yeast, and may be more similar to processes in higher eukaryotes. Full characterization of the N-glycosylation pathways in these organisms is very important, since protein glycosylation plays an integral role in processes such as pathogenicity and protein secretion.

The early steps of the asparagine-linked (N-linked) protein glycosylation pathways are similar in higher and lower eukaryotes (see reviews in Kornfeld and Kornfeld, *Ann. Rev. Biochem.* **54**:631-664, 1985; Moreman et al., *Glycobiol.* **4**:113-125, 1994; Herscovics and Orlean, *FASEB J.* 7:540-550, 1993; Herscovics, *Biochim.*

Biophys. Acta **1426**:275-285, 1999). Initially, an oligosaccharide precursor consisting of three glucose, nine mannose, and two N-acetylglucosamine molecules ($Glc_3Man_9GlcNAc_2$) is co-translationally transferred to the newly synthesized polypeptide in the endoplasmic reticulum (ER). In the ER, α-glucosidases I and II first remove the three glucose molecules. In the ER and/or Golgi apparatus, α-mannosidase(s) then remove one or more of the mannose residues from the precursor. In higher eukaryotes (i.e., mammals), α-1,2-mannosidases remove a total of four mannose residues, yielding $Man_5GlcNAc_2$ which is the precursor for complex, hybrid, and high-mannose N-glycans. In the yeast *S. cerevisiae*, however, an ER-specific mannosidase, α-1,2 mannosidase, removes only a single mannose residue, producing $Man_8GlcNAc_2$ (Herscovics, *Biochim. Biophys. Acta* **1426**:275-285, 1999).

Subsequent steps in the pathways in higher and lower eukaryotes are quite different. In higher eukaryotes, following the addition of a single GlcNAc to $Man_5GlcNAc_2$ by GlcNAc transferase I (GnT I), mannosidase II removes two additional mannose groups, producing $GlcNAcMan_3GlcNAc_2$. Various transferases, such as GnT II, fucosyl transferase, galactosyl transferase, and sialyl transferase, assemble the oligosaccharide into its final structure. In higher eukaryotes a variety of different carbohydrate units can thus be attached to a common precursor to form an array of distinct N-glycans. In *S. cerevisiae*, after the removal of a single mannose, various mannosyltransferases then add additional mannose units to $Man_8GlcNAc_2$ to form large high-mannose N-glycans containing up to 13 mannose units, and even larger mannan outer chains containing up to 200 mannose residues (Herscovics and Orlean, *FASEB J.* 7:540-550, 1993). The removal of mannose residues from the glycan chain during the initial stages of processing appears to be significantly different for higher and lower eukaryotes, resulting in quite different N-glycan structures for these organisms. Based on the evolutionary history of filamentous fungi, their N-glycan processing is likely to contain elements of the pathways in both mammals and yeast. Filamentous fungi do not produce complex N-glycans because such fungi lack the further processive transferases; however, the initial oligosaccharide precursor is trimmed to $Man_5GlcNAc_2$.

The α-mannosidases have been classified previously into two independently derived groups, Class 1 and Class 2, based on biochemical properties, substrate specificity, inhibitor profiles, and sequence alignments, (Daniel et al., *Glycobiol.* **4**:551-566, 1994; Moreman et al., *Glycobiol.* **4**:113-125, 1994; Eades et al., *Glycobiol.* **8**:17-33, 1998). The first group contains the α-1,2-mannosidases found in the ER and the Golgi apparatus, including the ER Man_9-mannosidase, ER endomannosidase, and the Golgi mannosidase I. The second group of α-mannosidases is more heterogeneous and contains the lysosomal mannosidases, the Golgi mannosidase II, and a distantly related group of enzymes, including the rat ER/cytosolic mannosidase (Bischoff et al., *J. Biological Chem.* **28**:17110-17117, 1990), yeast vacuolar mannosidase (Yoshihisa and Anraku, *Biochem. Biophys. Res. Comm.* **163**:908-915, 1989), and the *A. nidulans* Class 2 mannosidase (Eades et al., *Glycobiol.* **8**:17-33, 1998).

Summary of the Invention

The invention stems from the discovery of the amino acid sequences and corresponding nucleic acid sequences for three mannosidases. These mannosidases are particularly useful for altering the glycosylation patterns of macromolecules such as proteins. Glycosylation affects many properties of a glycoprotein, including protein folding, protease resistance, intercellular trafficking, compartmentalization, secretion, inter- and intra-molecular associations, intermolecular affinities, tissue targeting and biological half-life. Glycosylation patterns may also significantly alter the biological activity, solubility, clearance, intermolecular aggregation, and antigenicity, especially for those proteins having therapeutic utility. Thus, the present invention enables glycoproteins to be engineered to be more effectively used and produced.

One aspect of the present invention provides the amino acid sequences of three mannosidase proteins. The invention also provides variants of the disclosed mannosidase proteins. These variants can differ from the disclosed sequences by one or more conservative amino acid substitutions. Additionally, the invention provides variants of the disclosed mannosidase proteins having at least 60% sequence identity to the disclosed mannosidase amino acid sequences.

According to another aspect of the invention, mannosidase-specific binding agents are provided. These specific binding agents bind substantially only the mannosidases and mannosidase variants described above.

According to another aspect of the invention, respective nucleic acid sequences are provided that encode the mannosidases summarized above. These nucleic acid sequences can be operably linked to control sequences and incorporated into any of various vectors. The resulting recombinant vectors are useful for transforming any of various host cells. Once transformed, the host cell can produce the mannosidases of the present invention. Host cells can be obtained from fungi, plants, bacteria, animals, yeast, and insects.

According to yet another aspect of the invention, methods are provided for altering the glycosylation pattern of target proteins using one or more of the mannosidases of the present invention. The mannosidase can be placed in contact with a protein and allowed to alter the glycosylation pattern of the protein. The methods can be practiced either *in vivo* by creating a transgenic host cell that over-expresses or under-expresses one or more of the mannosidases, or *in vitro* by first isolating one or more of the mannosidases of the present invention, and then contacting the isolated mannosidase with a target protein.

According to yet another aspect of the invention, isolated nucleic acid molecules and amino acid molecules are provided that have, by way of example, at least 15, 20, 30, 40, or 50 contiguous nucleotides or amino acid residues with the sequences shown in SEQ ID NOS: 1, 4, and 17; or the complementary strands thereorof, or SEQ ID NOS: 3, 6, or 18, respectively.

Brief Description of the Drawings

Figures 1A – 1J show a DNA sequence of *A. nidulans* mannosidase 1A. The derived amino acid sequence is indicated by single-letter designations. The hydrophobic transmembrane region is indicated by a dotted underline. Upstream elements are underlined. The intron is indicated in bold type and consensus splice sites are underlined. The stop codon is indicated by a triple asterisk (***).

Figures 2A-2G show a DNA sequence of *A. nidulans* mannosidase 1B. The derived amino acid sequence is indicated by single-letter designations. Upstream

elements are underlined. The introns are indicated in bold type and consensus splice sites are underlined. The stop codon is indicated by a triple asterisk (***).

Figure 3 shows a Kyte-Doolittle hydropathy plot of the predicted amino acid sequences of (A) *A. nidulans* mannosidase 1A and (B) *A. nidulans* mannosidase 1B. The vertical axis is the hydrophobicity of a given region of the protein (Kyte and Doolittle, *J. Mol. Biol.* **157**:105-132, 1982) with positive values representing hydrophobic regions of the protein, and negative values representing hydrophilic regions of the protein.

Figure 4 shows a sequence-similarity matrix generated from the ClustalW (Thompson et al., *Nucl. Acids Research* **22**:4673-4680, 1994) sequence alignment of 13 Class 1 α-mannosidases and the pairwise percent similarity of the amino acid sequences for all possible pairs of sequences.

Figure 5 shows a dendrogram showing sequence relationships of mannosidases. The dendrogram was generated from the ClustalW sequence alignment. Three major groups of related enzymes are shown in shaded boxes.

Figures 6A and 6B show the amino acid and nucleic acid sequences of mannosidase 1C.

Sequence Listings

The nucleic and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three-letter code for amino acids. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand.

SEQ ID NO: 1 is the nucleic acid sequence of the mannosidase 1A gene.

SEQ ID NO: 2 is the nucleic acid sequence of the mannosidase 1A cDNA.

SEQ ID NO: 3 is the deduced amino acid sequence of mannosidase 1A.

SEQ ID NO: 4 is the nucleic acid sequence of the mannosidase 1B gene.

SEQ ID NO: 5 is the nucleic acid sequence of the mannosidase 1B cDNA.

SEQ ID NO: 6 is the deduced amino acid sequence of mannosidase 1B.

SEQ ID NO: 7 is a 5'-splice site sequence of the mannosidase 1A gene.

SEQ ID NO: 8 is a 5'-splice site consensus sequence of filamentous fungi.

SEQ ID NO: 9 is an internal lariat sequence of the mannosidase 1A gene.

SEQ ID NO: 10 is an internal lariat consensus sequence of filamentous fungi.

SEQ ID NO: 11 is a 5'-splice site sequence of the mannosidase 1B gene.

SEQ ID NO: 12 is an internal lariat sequence of the mannosidase 1B gene.

SEQ ID NO: 13 is the amino acid sequence used to generate a forward primer.

SEQ ID NO: 14 is the amino acid sequence used to generate a reverse primer.

SEQ ID NO: 15 is a PCR primer useful for identifying mannosidases.

SEQ ID NO: 16 is a PCR primer useful for identifying mannosidases.

SEQ ID NO: 17 is the nucleic acid sequence of the mannosidase 1C gene.

SEQ ID NO: 18 is the deduced amino acid sequence of mannosidase 1C.

SEQ ID NO: 19 is the deduced amino acid sequence of a consensus splice site.

Detailed Description

I. Definitions

"Sequence Identity." The similarity between two nucleic acid sequences or between two amino acid sequences is expressed in terms of the level of sequence identity shared between the sequences. Sequence identity is typically expressed in terms of percentage identity; the higher the percentage, the more similar the two sequences are.

Methods for aligning sequences for comparison purposes are well known in the art. Various programs and alignment algorithms are described in: Smith & Waterman, *Adv. Appl. Math.* **2:**482, 1981; Needleman & Wunsch, *J. Mol. Biol.* **48:**443, 1970; Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* **85:**2444, 1988; Higgins & Sharp, *Gene* **73:**237-244, 1988; Higgins & Sharp, *CABIOS* **5:**151-153, 1989; Corpet et al., *Nucleic Acids Research* **16:**10881-10890, 1988; Huang, et al., *Computer Applications in the Biosciences* **8:**155-165, 1992; and Pearson et al., *Methods in Molecular Biology* **24:**307-331, 1994. Altschul et al., *J. Mol. Biol.*, **215:**403-410, 1990, presents a detailed consideration of sequence alignment methods and homology calculations.

The NCBI Basic Local Alignment Search Tool (BLAST™, Altschul et al.. *J. Mol. Biol.*, **215**:403-410, 1990) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence-analysis programs blastp, blastn, blastx, tblastn and tblastx. BLAST™ can be accessed on the NCBI website. A description of how to determine sequence identity using this program is available on the internet at the NCBI website.

For comparisons of amino acid sequences of greater than about 30 amino acids, the "Blast 2 sequences" function in the BLAST™ program is employed using the default BLOSUM62 matrix set to default parameters, (gap existence cost of 11, and a per-residue gap cost of 1). When aligning short peptides (fewer than about 30 amino acids), the alignment should be performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties). Proteins with even greater similarity to the reference sequences will show increasing percentage identities when assessed by this method, such as at least 45%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% sequence identity.

"Substantial similarity" A first nucleic acid is "substantially similar" to a second nucleic acid if, when optimally aligned (with appropriate nucleotide insertions or deletions) with the other nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 60%, 75%, 80%, 85%, 90% or 95% of the nucleotide bases. Sequence similarity can be determined by comparing the nucleotide sequences of two nucleic acids using the BLAST™ sequence analysis software (blastn) available from The National Center for Biotechnology Information. Such comparisons may be made using the software set to default settings (expect = 10, filter = default, descriptions = 500 pairwise, alignments = 500, alignment view = standard, gap existence cost = 11, per residue existence = 1, per residue gap cost = 0.85). Similarly, a first polypeptide is substantially similar to a second polypeptide if it shows sequence identity of at least about 75%-90% or greater when optically aligned and compared using BLAST™ software (blastp) using default settings.

"Specific Binding Agent." A "specific binding agent" is an agent that is capable of specifically binding to the mannosidases of the present invention, and may include polyclonal antibodies, monoclonal antibodies (including humanized monoclonal antibodies) and fragments of monoclonal antibodies such as Fab, F(ab')2 and Fv fragments, as well as any other agent capable of specifically binding to the epitopes on the proteins.

"Operably linked." A first nucleic acid sequence is "operably linked" with a second nucleic acid sequence whenever the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein-coding regions, in the same reading frame.

"Isolated." An "isolated" biological component (such as a nucleic acid or protein or organelle) is a component that has been substantially separated or purified away from other biological components in the cell of the organism in which the component naturally occurs, i.e., other chromosomal and extra-chromosomal DNA, RNA, proteins, and organelles. Nucleic acids and proteins that have been "isolated" include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids and proteins prepared by recombinant expression in a host cell, as well as chemically synthesized nucleic acids.

"Recombinant." A "recombinant" nucleic acid is one having a sequence that is not naturally occurring or has a sequence made by an artificial combination of two otherwise-separated, shorter sequences. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques.

"Transformed." A "transformed" cell is a cell into which a nucleic acid molecule has been introduced by molecular biology techniques. As used herein, the term "transformation" encompasses all techniques by which a nucleic acid molecule might be introduced into such a cell, including transfection with a viral vector, transformation with a plasmid vector, and introduction of naked DNA by electroporation, lipofection, and particle gun acceleration.

"Vector." A "vector" is a nucleic acid molecule as introduced into a host cell, thereby producing a transformed host cell. A vector may include nucleic acid sequences, such as an origin of replication, that permit the vector to replicate in a host cell. A vector may also include one or more selectable marker genes and other genetic elements known in the art.

"Glycosylation pattern." The "glycosylation pattern" is the characteristic structure (including branch structure), number, or location of oligosaccharide structures associated with a macromolecule, such as a protein.

"DNA construct." The term "DNA construct" is intended to denote any nucleic acid molecule of cDNA, genomic DNA, synthetic DNA or RNA origin. The term "construct" is intended to denote a nucleic acid segment that may be single- or double-stranded, and that may be based on a complete or partial naturally occurring nucleotide sequence encoding one or more of the mannosidase genes of the present invention. It is understood that such nucleotide sequences include intentionally manipulated nucleotide sequences, e.g., subjected to site-directed mutagenesis, and sequences that are degenerate as a result of the genetic code. All degenerate nucleotide sequences are included within the scope of the invention so long as the mannosidase enzyme encoded by the nucleotide sequence maintains the ability to hydrolytically remove terminal mannoside residues.

"Mannosidase activity." The phrase "mannosidase activity" describes the enzymatic catalysis of the hydrolytic removal of terminal mannoside residues.

"Probes and primers." Nucleic acid probes and primers may be readily prepared based on the nucleic acid sequences provided by this invention. A "probe" comprises an isolated nucleic acid sequence attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed, e.g., in Sambrook et al. (eds.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel et al. (ed.) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New York (with periodic updates), 1987.

"Primers" are short nucleic acids, preferably DNA oligonucleotides 15 nucleotides or more in length, that are annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, then extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR) or other nucleic-acid amplification methods known in the art.

As noted, probes and primers are preferably 15 nucleotides or more in length, but, to enhance specificity, probes and primers of 20 or more nucleotides may be preferred.

Methods for preparing and using probes and primers are described, for example, in Sambrook et al. (ed.), *Molecular Cloning: A Laboratory Manual,* 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; Ausubel et al. (ed.), *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New York (with periodic updates), 1987; and Innis et al., *PCR Protocols: A Guide to Methods and Applications*, Academic Press: San Diego, 1990. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge, MA). One of skill in the art will appreciate that the specificity of a particular probe or primer increases with the length of the probe or primer. For example, a primer comprising 20 consecutive nucleotides will anneal to a target with a higher specificity than a

corresponding primer of only 15 nucleotides. Thus, in order to obtain greater specificity, probes and primers may be selected that comprise, by way of example, 10, 20, 25, 30, 35, 40, 50 or more consecutive nucleotides.

II. Methods of Producing Mannosidase 1A, 1B, and 1C

A. Cloning Nucleic Acid Sequences Encoding Mannosidase

Provided with the nucleic acid sequences of the genes encoding the mannosidases 1A, 1B, and 1C (SEQ ID NOS: 1, 4, and 17, respectively), one of ordinary skill in the art will appreciate that several different methods can be used to isolate the genes and the cDNAs encoding the corresponding mannosidases. One example of such a method is the polymerase chain reaction (PCR) (U.S. Pat. No. 4,683,202 to Mullis; and Saiki et al., *Science* **239:**487-491, 1988).

When using PCR to isolate a sequence encoding the gene, a primer can be designed that targets the extreme 5' end of the sequence, and a second primer can be designed that targets the extreme 3' end of the sequence. For example the 5' primer (5'-GGYGGYCTNGGYGARTCNTTCTACGAGTA-3'; SEQ ID NO: 15) and the 3' primer (5'-GTANAGGTACTTNAGNGTCTCNGCNAGRHAGAA-3'; SEQ ID NO: 16) are used in a PCR (polymerase chain reaction) procedure to generate multiple copies of the gene. The copies are isolated by separation on an agarose gel. The fragment of interest is then removed from the gel, and ligated into an appropriate vector.

Alternatively, the gene can be created by engineering synthetic strands of DNA that partially overlap each other (Beaucage & Caruthers, *Tetrahedron Letters* **22:**1859-1869, 1981; Matthes et al., *Embo. J.* **3:**801-805, 1984). The synthetic strands are annealed and a DNA polymerase is used to fill in the single-stranded regions. The resulting synthetic double-stranded DNA molecule can be cloned into a vector.

For use as primers and probes, nucleic acid sequences can contain at least 15 contiguous nucleic acid molecules of either of the sequences shown in Seq. ID. No. 1 and SEQ ID NO: 4, or either of the complementary strands of the molecules shown in SEQ ID NO: 1 and SEQ ID NO: 4. The nucleic acid sequences are useful

for performing hybridization protocols, such as Northern blots or Southern blots as described in Sambrook et al., (eds.), *Molecular Cloning, A Laboratory Manual*, 2d ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

These hybridization protocols can be used to identify nucleic acid sequences that are substantially similar to those shown in SEQ ID NOS: 1, 4, or 17. A successful hybridization to such sequences indicates that the analogous nucleic acid sequence hybridizes to the oligonucleotide probe that comprises at least a fragment of the sequences shown in SEQ ID NOS: 1, 4, or 17. Generally hybridization conditions are classified into categories, for example very high stringency, high stringency, and low stringency. The conditions corresponding to these categories are provided below.

Very High Stringency (detects sequences that share 90% sequence identity)

Hybridization in	5x	SSC	at	65°C	16 hours
Wash twice in	2x	SSC	at	room temp.	15 minutes each
Wash twice in	0.2x	SSC	at	65°C	20 minutes each

High Stringency (detects sequences that share 80% sequence identity or greater)

Hybridization in	3x	SSC	at	65°C	16 hours
Wash twice in	2x	SSC	at	room temp.	15 minutes each
Wash twice in	0.5x	SSC	at	55°C	20 minutes each

Low Stringency (detects sequences that share greater than 50% sequence identity)

Hybridization in	3x	SSC	at	65°C	16 hours
Wash twice in	2x	SSC	at	room temp.	20 minutes each

Mannosidase-encoding nucleic acid sequences according to the invention also encompass mannosidase enzymes that differ in amino acid sequence from the mannosidase 1A, 1B, and 1C sequences of SEQ ID NOS: 1, 4,and 17, and that maintain mannosidase activity. Such proteins may be produced by changing the

cDNA nucleotide sequence of mannosidase 1A, 1B, or 1C, by changing the sequence of the respective genes using standard procedures such as site-directed mutagenesis, or by performing the polymerase chain reaction. The simplest modifications involve substituting one or more amino acids with other amino acids having similar biochemical properties. These so-called "conservative substitutions" usually have minimal impact on the activity of the resultant protein. Table 1 shows amino acids that may be substituted for an original amino acid in a protein and that are regarded as conservative substitutions.

Table 1

Original Residue	Conservative Substitutions
ala	ser
arg	lys
asn	gln; his
asp	glu
cys	ser
gln	asn
glu	asp
gly	pro
his	asn; gln
ile	leu; val
leu	ile; val
lys	arg; gln; glu
met	leu; ile
phe	met; leu; tyr
ser	thr
thr	ser
trp	tyr
tyr	trp; phe
val	ile; leu

More substantial changes in enzymatic function or other features may be obtained by selecting substitutions that are less conservative than those in Table 1, i.e., selecting residues that differ more significantly in their effect on maintaining: (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Substitutions that

generally produce the greatest changes in protein properties are those in which: (a) a hydrophilic residue, e.g., seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g., leucyl, isoleucyl, phenylalanyl, valyl, or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histadyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine. The effects of these amino acid substitutions or deletions or additions may be assessed for mannosidase protein derivatives by analyzing the ability of the respective modified polypeptide to catalyze the removal of terminal mannoside residues.

Variant mannosidase cDNA or genes may be produced by standard DNA mutagenesis techniques, for example, M13 primer mutagenesis. Details of these techniques are provided in Sambrook et al. (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Ch. 15, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, and Ausubel et al. (ed.) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New York (with periodic updates), 1987. By the use of such techniques, variants may be created that differ slightly from the mannosidase cDNA or gene sequences specifically disclosed, yet that still encode a protein having mannosidase activity. DNA molecules and nucleotide sequences that are derivatives of those specifically disclosed herein and that differ from those disclosed by the deletion, addition, or substitution of nucleotides while still encoding a protein having mannosidase activity are comprehended by this invention. In their simplest form, such variants may differ from the disclosed sequences by alteration of the coding region to fit the codon usage bias of the particular organism into which the molecule is to be introduced.

Alternatively, the coding region may be altered by taking advantage of the degeneracy of the genetic code to alter the coding sequence in such a way that, while the nucleotide sequence is substantially altered, it nevertheless encodes a protein having an amino acid sequence identical or substantially similar to the mannosidase 1A and 1B sequences specifically disclosed in SEQ ID NOS: 3 and 6. For example, the fourth amino acid residue of the mannosidase 1A gene is alanine and is encoded

in the mannosidase 1A open reading frame (ORF) by the nucleotide codon triplet GCA. Because of the degeneracy of the genetic code, three other nucleotide codon triplets-GCT, GCC and GCG-also code for alanine. Thus, the nucleotide sequence of the mannosidase 1A ORF could be changed at this position to any of these three codons without affecting the amino acid composition of the encoded protein or the characteristics of the protein. Based upon the degeneracy of the genetic code, variant DNA molecules may be derived from the cDNA and gene sequences disclosed herein using standard DNA mutagenesis techniques as described above, or by synthesis of modified DNA sequences. Thus, this invention also encompasses nucleic acid sequences that encode either a mannosidase 1A or 1B protein but that vary from the disclosed nucleic acid sequences due to the degeneracy of the genetic code.

B. Vectors

The choice of expression vector depends in part on the type of host cell that will be used for the expression of the mannosidase enzyme. The promoter sequence used must be recognizable by the enzymes responsible for translation within the host cell. One of skill in the art will appreciate that there are a number of regulatory sequences known in the art that function in conjunction with bacterial host cells, plant host cells, insect host cells, yeast host cells, and fungal host cells.

Examples of constitutive plant promoters that may be useful for expressing the cDNA include: the cauliflower mosaic virus (CaMV) 35S promoter, which confers constitutive, high-level expression in most plant tissues (see, e.g., Odel et al., *Nature* **313**:810, 1985), nopaline synthase promoter (An *et al.*, *Plant Physiol.* **88**:547, 1988); and the octopine synthase promoter (Fromm et al., *Plant Cell* **1**:977, 1989).

Examples of suitable promoters for use in filamentous fungus host cells are, for instance, the ADH3 promoter (McKnight et al., *Embo. J.* **4**:2093-2099, 1985) or the tpiA promoter (Blattner et al., *Science* **27**:1453-1474, 1997). Examples of other useful promoters are those derived from the genes encoding *A. oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *A. niger* neutral α-amylase, *A. niger* acid-stable α-amylase, *A. niger* or *A. awamori* glucoamylase (gluA),

Rhizomucor miehei lipase, *A. oryzae* alkaline protease, *A. oryzae* triose phosphate isomerase, or *A. nidulans* acetamidase. It is also possible that the native mannosidase promoter could be used to express the enzyme. Similarly, one of ordinary skill in the art will appreciate that there are several promoters available that can be used in insect expression systems (e.g., baculovirus) and bacterial expression systems.

To direct the enzyme into the secretory pathway of the host cell, a secretory signal sequence (also known as a leader sequence, "prepro" sequence or "pre" sequence) may be provided in the vector. The secretory signal sequence is joined to the DNA sequence encoding the enzyme in the correct reading frame. Secretory signal sequences are commonly positioned 5' to the DNA sequence encoding the enzyme. The secretory signal sequence may be the sequences normally associated with the enzyme or from a gene encoding another secreted protein.

The vector will also most likely contain a selectable marker. The selectable marker allows host cells that have been successfully transformed with the construct of interest to be identified. Selectable markers for use in plant cells are, for example Bosta^r, Kan^r, and various other herbicide resistance genes. Selectable markers for use in mammalian cells are, for example, Amp^r and Kan^r. Selectable markers for use in filamentous fungi include, for example, amdS, pyrG, argB, niaD and sC.

C. Transformation

The DNA construct of the invention may be either homologous or heterologous to the host in question. If homologous to the host cell, i.e., produced by the host cell in nature, the construct typically will be operably connected to another promoter sequence or, if applicable, another secretory signal sequence and/or terminator sequence than in its natural environment. In this context, the term "homologous" is intended to include a cDNA sequence encoding the mannosidase 1A; 1B, or 1C native to the host *A. nidulans*. The term "heterologous" is intended to include a DNA sequence not expressed by the host cell in nature. Thus, the DNA sequence may be from another organism, or it may be a synthetic sequence.

The host cell of the invention, into which the DNA construct or the recombinant expression vector of the invention is to be introduced, may be any cell

capable of producing the disclosed mannosidase enzymes. Such cells include bacteria cells, yeast cells, fungal cells, insect cells, and higher eukaryotic cells.

Various methods of introducing the DNA construct into host cells are well known in the art. For example, in some species, the Ti plasmid of *A. tumefaciens* can be used to transform host cells (Gouka et al., *Nature Biotechnology* **6**:598-602, 1999), the host cell can also be transformed using gene blasting techniques and standard chemical treatments.

D. Production

The production of recombinant mannosidase 1A, 1B, or 1C may be accomplished using any of various transformed hosts, such as bacteria, plants, mammalian cell culture, whole mammals, insects, yeast, and fungi. As mentioned above, the selection of suitable promoter sequences and other regulatory elements will be specific to the host organism that will be used to produce the protein. Additionally, within each group of potential hosts there are several species that can potentially be used to produce the protein. For example, filamentous fungi, e.g., *Aspergillus* spp., *Neurospora* spp., *Fusarium* spp. or *Trichoderma* spp., can be used to produce mannosidase. More specifically, individual strains of *Aspergillus*, such as *A. oryzae, A. nidulans*, or *A. niger*, can be used. The use of *Aspergillus* spp. for the expression of proteins is described in, e.g., EP 272,277 and EP 230,023. Transformation techniques specific for the transformation of *F. oxysporum* have been described by Malardier et al., *Gene* **78**:147-156, 1988.

Various yeast strains and yeast-derived vectors are commonly used for the expression of heterologous proteins. For instance, *Pichia pastoris* expression systems, obtained from Invitrogen (San Diego, California), may be used to practice the present invention. Such systems include suitable *Pichia pastoris* strains, vectors, reagents, transformants, sequencing primers, and media. Available strains include KM71H a prototrophic strain, SMD1168H a prototrophic strain, and SMD1168 a pep4 mutant strain (Invitrogen Product Catalogue, 1998, Invitrogen, Carlsbad CA).

Non-yeast eukaryotic vectors may be used with equal facility for expression of proteins encoded by modified nucleotides according to the invention. Mammalian vector/host cell systems containing genetic and cellular control

elements capable of carrying out transcription, translation, and post-translational modification are well known in the art. Examples of such systems are the well-known Baculovirus system, the ecdysone-inducible mammalian expression system that uses regulatory elements from *Drosophila melanogaster* to allow control of gene expression, and the Sindbis viral-expression system that allows high-level expression in a variety of mammalian cell lines, all of which are available from Invitrogen.

The cloned expression vector may be transformed into any of various cell types for expression of the cloned nucleotide. Many different types of cells may be used to express modified nucleic acid molecules. Examples include cells of yeasts, fungi, insects, mammals, and plants, including transformed and non-transformed cells. For instance, common mammalian cells that could be used for the invention include HeLa cells, SW-527 cells (ATCC deposit #7940), WISH cells (ATCC deposit #CCL-25), Daudi cells (ATCC deposit #CCL-213), Mandin-Darby bovine kidney cells (ATCC deposit #CCL-22), and Chinese hamster ovary (CHO) cells (ATCC deposit #CRL-2092). Common yeast cells include *Pichia pastoris* (ATCC deposit #201178) and *Saccharomyces cerevisiae* (ATCC deposit #46024). Insect cells include cells from *Drosophila melanogaster* (ATCC deposit #CRL-10191), the cotton bollworm (ATCC deposit #CRL-9281) and from *Trichoplusia ni* egg cell homoflagellates. Fish cells that may be used include those from rainbow trout (ATCC deposit #CLL-55), salmon (ATCC deposit #CRL-1681), and zebrafish (ATCC deposit #CRL-2147). Amphibian cells that may be used include those of the bullfrog, *Rana catesbelana* (ATCC deposit #CLL-41). Reptile cells that may be used include those from Russell's viper (ATCC deposit #CCL-140). Plant cells that could be used include *Chlamydomonas* cells (ATCC deposit #30485), *Arabidopsis* cells (ATCC deposit #54069) and tomato plant cells (ATCC deposit #54003). Many of these cell types are commonly used and are available from the ATCC as well as from commercial suppliers such as Pharmacia (Uppsala, Sweden), and Invitrogen (San Diego, California).

Expressed protein may be accumulated within a cell or may be secreted from the cell. Such expressed protein may then be collected and purified. This protein

may then be characterized for activity and heat stability and may be used to practice any of the various methods according to the invention.

E. Isolation and Purification of Mannosidase 1A, 1B, and 1C

The mannosidases of the present invention can be isolated and purified from either transgenic host cells, or from wild-type host cells. The purification of the mannosidase enzymes of the present invention is achieved by first isolating the enzymes from the other cellular components and then purifying the enzymes.

1. Separation of Enzymes from Cellular Components

A variety of methods are known in the art for separating enzymes from other cellular components (complexes). Such methods can involve treatment with harsh chemicals and/or severe temperatures. Typically, protein complexes can be disrupted through the use of reducing agents, denaturants, freeze/thaw cycles, mechanical shearing, decompression/compression, sonication, agitation, and/or increased temperatures. Protein complexes also can be disrupted using a combination of such techniques.

Reducing agents are capable of donating hydrogen atoms, and thus serve to cleave disulfide bonds. Reducing agents particularly disrupt disulfide bonds that link two proteins or portions of the same protein together. Commonly used reducing regents are β-mercaptoethanol, dithiothreitol (DTT), and trialkyl phosphines. The ability of a reducing agent to disrupt a complex is increased by increasing the temperature at which a mixture of the reducing agent and complex is incubated.

Denaturants serve to relax the conformational structure of proteins. Any of various denaturants can be employed to separate a protein from other cellular components. Examples of denaturants useful in conjunction with ion-exchange columns are urea and formamide. However, denaturants such as guanidine hydrochloride or guanidine thiocyanate may be useful for separating proteins from other cellular components in methods not involving an ion-exchange column. The ability of a particular denaturant to relax a protein is enhanced by increasing the temperature at which the sample is incubated.

2. Purification of Enzymes from Cellular Components

Purification of enzymes from a crude lysate can be achieved by successive rounds of chromatography. In between each round of chromatography, the sample is assayed for enzymatic activity. Examples 7 and 8, below, describe activity assays that can be used to monitor the recovery of mannosidase activity, respectively. Typically, purification involves a multi-step procedure that includes well-known chromatographic techniques (Robyt and White, *Biochemical Techniques Theory and Practice*, Waveland Press, Inc., 1990). In such multi-step procedures, the enzymes are separated by exploiting their different physical characteristics. The following discussion provides a broad description of various chromatography techniques that can be used in either a single-step process or in a multi-step process. A single type of chromatography can be repeatedly used (rather than changing the chromatography step each time) to purify the enzymes of the present invention.

i. Column Chromatography

One method of isolating an enzyme according to the present invention is adsorption chromatography. This method exploits a protein's differential affinity for the medium in a column, compared to the protein's affinity for the eluting solvent. An example of a suitable medium for use in the column is hydroxyapatite (crystalline calcium phosphate). Hydroxyapatite tends to adsorb acidic proteins that can be subsequently eluted with phosphate ions (phosphate ions have a high affinity for the calcium ions present in the hydroxyapatite).

Ion-exchange chromatography (a variation of adsorption chromatography) can also be used to isolate enzymes according to the present invention. In ion-exchange chromatography, a solid adsorbent is used that has charged groups chemically linked to an inert solid. An ionic charge on an enzyme molecule causes the molecule to attach to an oppositely charged group on the solid support. The enzyme is subsequently released from the support by passing a solution containing an ion gradient over the solid adsorbent. Examples of solid supports are DEAE-cellulose, DEAE-Sephadex™ , DEAE-Bio-Gel™, DEAE-Sepharose™, DEAE-Sephacryl™, DEAE-Trisacryl™, Q-Sepharose™, ecteola cellulose, QAE-cellulose,

express ion exchanger Q, PEI-cellulose, and other polystyrene-based anion exchangers (most of these support materials are available from Pharmacia).

Another suitable type of adsorption chromatography is affinity chromatography. This method involves covalently linking to an inert solid support a ligand having a binding affinity for the subject enzyme. Commonly, the ligand is a specific binding agent that selectively binds the enzyme as the enzyme molecule contacts the solid support. Alternatively, the enzymes can be purified based upon their hydrophobicity. For example, alkyl chains can be linked to the inert support to supply sites for hydrophobic bonding interactions between the support and the enzyme. The eluting solvent contains a hydrophobic gradient.

High-performance liquid chromatography (HPLC) is yet another suitable method. All of the major classes of chromatographic separations are compatible with this method, for example: adsorption, liquid-liquid partition, ion exchange, exclusion, and affinity chromatography can be used in conjunction with HPLC. Additionally, HPLC allows for reverse-phase and ion-pair partition. Reverse-phase partition is a relatively quick elution technique using a non-polar stationary phase and a polar mobile phase. Ion-pair partition involves pairing a charged polar substance with its counter-ion to create a less polar species that then flows through the column.

ii. Electrophoresis

Electrophoresis is a well-established technique for the separation and analysis of mixtures by differential migration and separation of molecules in an electric field, based on differences in mobility of the molecule through a support. Many different forms of electrophoresis have been developed to permit the separation of different classes of compounds. These forms include paper and cellulose acetate electrophoresis, thin-layer electrophoresis, gel electrophoresis, immunoelectrophoresis, and isoelectric focusing. Paper electrophoresis operates best for the separation of protein molecules having a relatively low molecular weight, and gel electrophoresis is better for use in isolating enzymes with higher molecular weights. A variety of different matrices are available for forming gels, but matrices with a relatively small pore size are most suitable for the separation of

two proteins that have similar physical characteristics. Furthermore, proteins may be separated based upon their molecular size by conducting electrophoresis under dissociating conditions, for example using sodium dodecyl sulfate (SDS). SDS relaxes the protein conformation and masks the ionic charge of the protein, hereby leaving the relative length of the protein as the principal distinguishing characteristic for purposes of separation from other proteins.

Separation also can be achieved using immunoelectrophoresis, in which proteins are separated on a gel, based upon the relative charge-to-mass ratio and antigenicity of the proteins. The proteins are separated on a gel. An antibody, specific for the protein of interest, is added to a well created in the gel, and the antibody is allowed to diffuse through the gel. A precipitate forms in regions in which the antibody reacts with the protein.

Finally, the enzymes can be purified by isoelectric focusing (IEF). IEF is a type of electrophoresis in which the protein is placed on a substrate having a pH gradient. The protein moves under the influence of an applied electrical field until the protein reaches a zone in the pH gradient corresponding to the isoelectric point of the protein.

Specific examples of previously published mannosidase-purification techniques are available in Castellion and Bretthauer, *Biochem. J.*, **324**:951-956, 1997; and Eades et al., *Glycobiology* **8**:17-33, 1998.

III. Mannosidase Specific Antibodies

Antibodies to the mannosidase enzymes of the present invention may be useful for purification of the enzymes. The provision of the amino acid sequences of the mannosidase 1A, 1B, and 1C enzymes allows for the production of specific antibody-based binding agents to these enzymes.

A. Production of an Antibodies to Mannosidase 1A, 1B and/or 1C

Monoclonal or polyclonal antibodies may be produced to either of the mannosidase enzymes, portions of the enzymes, or variants thereof. Optimally, antibodies raised against epitopes on these antigens will specifically detect the enzyme. That is, antibodies raised against the enzymes recognize and bind the

enzymes, but do not substantially recognize or bind to other proteins. The determination that an antibody specifically binds to an antigen is made by any one of a number of standard immunoassay methods, for instance, Western blotting, Sambrook et al. (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

To determine that a given antibody preparation (such as a preparation produced in a mouse against mannosidase 1A) specifically detects a mannosidase by Western blotting, total cellular protein is extracted from fungal cells and electrophoresed on an SDS-polyacrylamide gel. The proteins are then transferred to a membrane (for example, nitrocellulose) by Western blotting, and the antibody preparation is incubated with the membrane. After washing the membrane to remove non-specifically bound antibodies, the presence of specifically bound antibodies is detected by the use of an anti-mouse antibody conjugated to an enzyme such as alkaline phosphatase; application of 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium results in the production of a dense blue compound by immuno-localized alkaline phosphatase.

Antibodies that specifically detect mannosidase can be shown, by this technique, to bind substantially only the mannosidase band (having a position on the gel determined by the molecular weight of the mannosidase). Non-specific binding of the antibody to other proteins may occur and may be detectable as a weaker signal on the Western blot (which can be quantified by automated radiography). The non-specific nature of this binding can be recognized by one skilled in the art by the weak signal obtained on the Western blot relative to the strong primary signal arising from the specific anti-mannosidase binding.

Antibodies that specifically bind to mannosidase belong to a class of molecules that are referred to herein as "specific binding agents." Specific binding agents that are capable of specifically binding to the mannosidases of the present invention may include polyclonal antibodies, monoclonal antibodies, and fragments of monoclonal antibodies such as Fab, $F(ab')_2$ and Fv fragments, as well as any other agent capable of specifically binding to one or more epitopes on the proteins.

Substantially pure mannosidase 1A or 1B suitable for use as an immunogen can be isolated from transfected cells, transformed cells, or from wild-type cells.

Concentration of protein in the final preparation is adjusted, for example, by concentration on an Amicon filter device (Charlotte, North Carolina), to the level of a few micrograms per milliliter. Alternatively, peptide fragments of either of the mannosidases may be utilized as immunogens. Such fragments may be chemically synthesized using standard methods, or may be obtained by cleavage of the whole mannosidase enzyme followed by purification of the desired peptide fragments. Peptides as short as three or four amino acids in length are immunogenic when presented to an immune system in the context of a Major Histocompatibility Complex (MHC) molecule, such as MHC class I or MHC class II. Accordingly, peptides comprising at least 3 and preferably at least 4, 5, 6 or more consecutive amino acids of the disclosed mannosidase amino acid sequences may be employed as immunogens for producing antibodies.

Because naturally occurring epitopes on proteins frequently comprise amino acid residues that are not adjacently arranged in the peptide when the peptide sequence is viewed as a linear molecule, it may be advantageous to utilize longer peptide fragments from the mannosidase amino acid sequences for producing antibodies. Thus, for example, peptides that comprise at least 10, 15, 20, 25, or 30 consecutive amino acid residues of the amino acid sequence may be employed. Monoclonal or polyclonal antibodies to the intact mannosidase, or peptide fragments thereof may be prepared as described below.

B. Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibody to any of various epitopes of the mannosidase enzymes that are identified and isolated as described herein, can be prepared from murine hybridomas according to the classical method of Kohler & Milstein, *Nature*, **256:**495, 1975, or a derivative method thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein over a period of a few weeks. The mouse is then sacrificed, and the antibody-producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess unfused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where

growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall, *Enzymol.* **70**:419, 1980, or a derivative method thereof. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Harlow & Lane, *Antibodies, A Laboratory Manual,* Cold Spring Harbor Laboratory, New York, 1988.

C. Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogenous epitopes of a single protein can be prepared by immunizing suitable animals with the expressed protein, which can be unmodified or modified, to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less immunogenic than other molecules and may require the use of carriers and an adjuvant. Also, host animals vary in response to site of inoculations and dose, with both inadequate or excessive doses of antigen resulting in low-titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appear to be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis et al., *J. Clin. Endocrinol. Metab.* **33**:988-991, 1971.

Booster injections can be given at regular intervals, and antiserum harvested when the antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony et al., *Handbook of Experimental Immunology,* Wier, D. (ed.), Chapter 19, Blackwell, 1973. A plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/mL of serum (about 12 μM). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher, *Manual of Clinical Immunology*, Chapter 42, 1980.

D. Antibodies Raised by Injection of cDNA

Antibodies may be raised against the mannosidases of the present invention by subcutaneous injection of a DNA vector that expresses the enzymes in laboratory animals, such as mice. Delivery of the recombinant vector into the animals may be achieved using a hand-held form of the Biolistic system (Sanford et al., *Particulate Sci. Technol.* **5**:27-37, 1987, as described by Tang et al., *Nature* (London) **356**:153-154, 1992). Expression vectors suitable for this purpose may include those that express the cDNA of the enzyme under the transcriptional control of either the human β-actin promoter or the cytomegalovirus (CMV) promoter. Methods of administering naked DNA to animals in a manner resulting in expression of the DNA in the body of the animal are well known and are described, for example, in U.S. Patent Nos. 5,620,896 ("DNA vaccines against rotavirus infections"); 5,643,578 ("Immunization by inoculation of DNA transcription unit"); and 5,593,972 ("Genetic immunization"), and references cited therein.

E. Antibody Fragments

Antibody fragments may be used in place of whole antibodies and may be readily expressed in prokaryotic host cells. Methods of making and using immunologically effective portions of monoclonal antibodies, also referred to as "antibody fragments," are well known and include those described in Better & Horowitz, *Methods Enzymol.*, **178**:476-496, 1989; Better et al., in Streilein et al., eds., *Advances in Gene Technology: The Molecular Biology of Immune Disease & the Immune response (ICSU Short Reports)*, **10**:105, 1990; Glockshuber et al. *Biochemistry* **29**:1362-1367, 1990; and U.S. Patent Nos. 5,648,237 ("Expression of Functional Antibody Fragments"), No. 4,946,778 ("Single Polypeptide Chain Binding Molecules"), and No. 5,455,030 ("Immunotherapy Using Single Chain Polypeptide Binding Molecules"), and references cited therein.

IV. Using Mannosidase

A. *In vitro*

The mannosidases of the present invention are useful for modifying the glycosylation pattern proteins *in vitro*. In particular, the enzymes of the present invention are useful for modifying target proteins. Target proteins are specific proteins that are the desired product of a process, for example, a glycoprotein that will be used as a therapeutic agent.

The *in vitro* modification of a target protein is accomplished by first obtaining one or more of the mannosidases of the present invention. The mannosidase may be purified to achieve specificity of action. The mannosidase is then placed in contact with the target protein, allowing for the modification of the glycosylation pattern of the target protein in a controlled, *in vitro* environment. *In vitro* methods of modifying the glycosylation pattern of a protein are well known in the art, and examples of such methods are provided in U.S. Patent No. 5,834,251 to Maras, et al., herein incorporated by reference.

B. *In vivo*

The glycosylation pattern of a protein, a broad class of proteins, or all glycosylated proteins in a host cell can be modified *in vivo* through the use of molecular biology techniques. The provision of the disclosed nucleic acid sequences allows for both the up-regulation of the disclosed mannosides as well as the down-regulation of the disclosed mannosidases.

Up regulation of the protein can be accomplished by transforming a cell with a vector containing one or more of the disclosed mannosidase genes under the control of a promoter. The promoter can be for example a constitutive promoter, a tissue specific promoter, or an inducible promoter. Thus, the RNA encoding the enzyme(s) is produced and translated into the protein. The increased level of mannosidase in the cell will then alter the glycosylation pattern of proteins produced by the cell.

Conversely, the production of mannosidase by a cell can be down-regulated by transforming the cell with a vector encoding an antisense molecule, or a catalytic

nucleic acid molecule that targets the mannosidase specific nucleic acid sequences. This technique will likely cause the cell to produce less mannosidase and therefore, glycoproteins from the cell will be processed to a lesser extent.

In contrast to the above-described *in vitro* methods, *in vivo* methods can be used to produce multiple different proteins in the same cell, all of which display altered glycosylation patterns.

V. Examples

Example 1

Identification of *Aspergillus nidulans* α-1, 2-mannosidase 1A. Four previously published Class 1 α-1,2-mannosidase protein sequences were aligned, and conserved blocks were identified for PCR primer design. Selection of sequences for primer design was based on high sequence identity between the four proteins, and low codon redundancy. A codon-preference table for *A. nidulans* was used to decrease the codon redundancy in the primers and to select codons more likely to be found in the genomic DNA. Two primers were selected that contained the least redundancy and the highest possible annealing temperature. Amplification of *A. nidulans* genomic DNA with these two primers yielded a 900-bp product. This PCR product was cloned and the DNA sequence was determined for several representative clones. The DNA sequence of clone pGEM42-9 was used to search the GenBank database for sequence homology (BLAST ™ search). Several high-scoring matches were found with other Class 1 α-1,2-mannosidases, confirming that a portion of the Class 1 α-1,2-mannosidase from *A. nidulans* had been amplified.

The 900-bp PCR fragment was used as a template for the PCR amplification of an α-1,2-mannosidase-specific radiolabeled probe for library screening. An EMBL-3 library of genomic sequences was screened and several positive plaques were re-screened to isolate a single lambda clone containing the entire α-1,2-mannosidase gene. Two *Bam*H1 subclones were identified by Southern hybridization, using the same PCR-derived probe, that contained the 5' and 3' ends of the gene. These subclones were sequenced by manual and automated sequencing, using a combination of primer walking with specifically designed primers and restriction enzyme subcloning using universal sequencing primers, yielding 2- to 5-

fold redundancy sequencing of the full-length gene and several thousand base pairs (bp) of flanking sequence (Fig. 1).

Example 2

Characterization of the α-1, 2-mannosidase 1A gene. A BLAST™ search of the α-1,2-mannosidase 1A gene revealed two open reading frames (ORFs) that contained significant homology to the Class 1 α-mannosidases. These ORFs were separated by a region of DNA containing several stop codons, which indicated an intron sequence. This intron sequence was verified by reverse-transcription PCR (RT-PCR) from total RNA and comparison to the genomic DNA. The amplification product derived from the RNA, using a PCR primer pair that flanked the putative intron, was 50 bp shorter than the similar PCR product derived from genomic DNA. Analysis of the putative intron revealed several typical sequence motifs which are characteristic of eukaryotic intron sequences. The 5'-splice site (5'-GTAAGT-3' (SEQ ID NO: 7)) fit the consensus sequence for filamentous fungi (5'-GTANGT-3' (SEQ ID NO: 8)), and the 3'-splice site (5'-TAG-3') fit the consensus 5'-YAG-3' (Ballance, *Molecular Industrial Mycology – Systems and Applications for Filamentious Fungi*, Leon and Berka (eds.), Dekker, Inc., 1991; Ballance, *Yeast* 2:229-236, 1986; Gurr et al., *Gene Structure in Eukarayotic Microbes*, Kinghorn (ed.), IRL Press, 1987). The intron also contained an internal lariat sequence (5'-GCTGAC-3' (SEQ ID NO: 9)), located 15 bp upstream of the 3'-splice site, which fit the consensus 5'-(G/A)CT(G/A)AC-3' (SEQ ID NO: 19).

The deduced amino acid sequence of the *A. nidulans* α-1,2-mannosidase 1A gene was aligned with other published α-mannosidases in order to identify potential introns that do not shift the reading frame and do not contain stop codons but do increase the size of the putative gene product. In addition to the previously confirmed intron, two other regions of the ORF that did not align with other published sequences seemed to be "extra" DNA sequences. These regions were represented by large "gaps" in the multiple sequence alignment for all of the other α-1,2-mannosidase genes used in the alignment. To determine whether these sequences represented introns, or encoded a polypeptide sequence novel to the *A. nidulans* α-mannosidase, primer pairs were designed that flanked these regions and

used in RT-PCR. The amplification products derived from the RNA (RT-PCR) were the same size as the amplification products derived from genomic DNA, indicating that there were no introns present in these regions. These regions did not contain consensus splice motifs, though sometimes introns were present with less-conserved splice sequences (Gurr et al., *Gene Structure in Eukarayotic Microbes*, Kinghorn (ed.), IRL Press, 1987).

The position of the translational start codon of the first ORF was determined through examination of the DNA sequence near the stop codon that defined the 5' end of the ORF. A potential in-frame start codon (ATG) occurred 42 bp after the beginning of the ORF, while the next in-frame ATG codon was located 423 bp downstream. The first ATG codon was thus a good candidate for the translational start codon. Translation originating at the first start codon would produce a protein product with an N-terminus larger than other fungal α-mannosidases, but similar in size to the N-termini of mammalian and insect Class 1 α-mannosidases.

Though no strong consensus sequence surrounds the translational start codon in filamentous fungi (Ballance, *Yeast* **2**:229-236, 1986), there is a preference (97%) for a purine at position -3. The putative start codon for this gene has a G at position -3, and thus conforms to this rule. There is a TATA-like element at position -47 and several pyrimidine-rich blocks within 100 bp upstream of the putative start codon. These pyrimidine rich regions are often found in fungal promoters and may influence the level of transcription (Ballance, *Yeast* **2**:229-236, 1986). The 5'-untranslated region does not contain a CAAT-box upstream of the TATA-like element, but this is not unusual for fungal promoters.

The α-1,2-mannosidase 1A protein contained several charged N-terminal amino acids representative of a typical signal sequence motif. A Kyte-Doolittle hydropathy plot showed that the signal sequence was followed by a highly hydrophobic region approximately 15-16 amino acids in length, likely encoding a transmembrane domain, while the rest of the protein was relatively hydrophilic (Fig. 3). This α-1,2-mannosidase 1A protein from *A. nidulans* likely forms a type II transmembrane protein, which is a characteristic of other Class 1 α-1,2-mannosidases.

Example 3

Identification of *A. nidulans* α-1,2-mannosidase 1B. The deduced coding region of the *A. nidulans* α-1,2-mannosidase 1A gene was used to search the *A. nidulans* EST Sequencing Project Database using the BLAST™ algorithm to determine if there were multiple β-1,2-mannosidase genes were expressed in this organism. A sequence tag, containing 200 bp of information, was identified that showed significant similarity, but not 100% identity, to the α-1,2-mannosidase 1A gene, and thus represented a separate gene. This sequence tag was used to design new PCR primers specific for the novel gene (α-1,2-mannosidase 1B). The primers were designed in regions that, based on multiple sequence alignments of published α-1,2-mannosidase sequences, would not be expected to be well conserved, and thus would be gene-specific. The 200-bp amplification product from these primers was radiolabeled and used to probe the *A. nidulans* genomic library to select the full-length gene.

A single lambda clone containing the full-length α-1,2-mannosidase 1B gene was isolated and subcloned into plasmid DNA. A single 5.6-kb *Bam*HI clone containing the full-length gene was identified by Southern analysis. A series of restriction enzyme resections of the 5.6-kb clone were derived and the sequence was determined by automated fluorescence sequencing using the universal priming sites of the cloning vector. The full-length gene and several hundred bases of flanking sequence were sequenced with 2- to 4-fold redundancy at each base position (Fig. 2).

Example 4

Characterization of α-1,2-mannosidase 1B. The genomic DNA sequence for the *A. nidulans* α-1,2-mannosidase 1B gene was searched for open reading frames that showed homology to other α-1,2-mannosidase genes. The gene contained three open reading frames with homology with Class 1 α-1,2-mannosidases separated by two regions that contained several stop codons and caused a shift in the reading frame containing α-1,2-mannosidase homology. The two regions that disrupted the open reading frame were analyzed to identify

potential introns. Comparison of the reading frame of the gene with the *A. satoi* and *P. citrinum* genes verified that these regions appeared to contain extra DNA sequences, since the alignment seems to be disrupted near these regions. These regions were searched for consensus intron-splice sites. The first potential intron did indeed contain 5' and 3' consensus splice sites as well as an internal lariat consensus sequence. The sequence 5'-GTACGT-3' (SEQ ID NO: 11) fit the filamentous fungal consensus for a 5'-splice site (5'-GTANGT-3'), and the sequence 5'-TAG-3' fit the consensus for the 3'-splice site (5'-YAG-3'). This intron also contained the internal lariat sequence 5'-ACTGAC-3' (SEQ ID NO: 12) located 11 bp upstream of the 3'-splice site. The second putative intron also contained a consensus 5'-splice site (5'-GTACGT-3'), a consensus 3'-splice site (5'-CAG-3'), and an internal lariat consensus (5'-ACTGAC-3') which was located 11 bp upstream of the 3'-splice site. The two putative introns were verified by RT-PCR. In both cases, the RNA-derived amplification product was smaller than the DNA-derived amplification product, and the size difference corresponded to the predicted size of each intron. Interestingly, the position of each intron in the coding region is correlated directly to two of the introns found in the *P. citrinum* α-1,2-mannosidase gene, although the intron sequences themselves did not appear to be conserved. The N-terminal region of the first ORF likely contained the translational start site, which would probably be the first methionine after the stop codon defining the beginning of the ORF. Comparison of the coding region of this ORF with the *A. satoi* and *P. citrinum* α-1,2-mannosidase genes showed that the putative translational start site correlated positionally with the start sites of these genes. This start codon contained a purine at the -3 position, as expected (Ballance, *Yeast* **2**:229-236, 1986). A TATA-like element was found at position -76 (TATAT), and the upstream region contained several CT-rich tracts in the sense strand. Seven copies of the sequence CTCC appeared in the 100-bp region upstream of the TATA-like element and could represent an important promoter element. Additionally, there was a 7-bp direct repeat located immediately upstream of the TATA-like element (CCTCAT). The N-terminus of this protein contained a typical signal sequence for insertion into the endoplasmic reticulum and, as seen in a Kyte-Doolittle hydropathy plot (Fig. 3), contained a 10-15 bp hydrophobic region which likely encoded a transmembrane

domain. This protein, like other Class 1 α-1,2-mannosidases, is likely a type-II transmembrane protein normally localized in the ER or Golgi apparatus.

Example 5

Identification of *A. nidulans* α-1,2-mannosidase 1C: A second lambda clone containing the full length α-1,2-mannosidase 1C gene was recovered and two non-overlapping *Bam*HI subclones (4 kb and 6 kb) were isolated, which together contained the gene and flanking regions. Again, the sequence across the *Bam*H1 subcloning junction was verified to eliminate the possiblity of missing sequence. The gene and several hundred bp of flanking region were fully sequenced (Accession #: AF233287).

Example 6

Comparision of α1,2-mannosidase 1A, 1B, and 1C sequences: The DNA sequences of the three Class I α-1,2-mannosidases were analyzed to determine the amino acid coding sequence, including determination of the correct reading frame, identification of potential intron sequences, and identification of the correct translational start codon for each gene. A BLAST search of theα-1,2-mannosidase IA gene revealed two open reading frames (ORFs) separated by a region of DNA containing several stop codons which could indicate the presence of an intron sequence. To verify the presence of an intron at this site, PCR products spanning the putative intron were amplified from reverse-transcribed RNA (RT-PCR), cloned into vector and sequenced. Comparison of the α-mannosidase IA sequence with RT-PCR sequence verified the presence of a 50 bp intron at the expected splice junction. The intron contained a 5'-splice site (5'-GTAAGT-3') which matched the consensus sequence for filamentous fungi (5'-GTANGT-3'), and a 3'-splice site (5'-TAG-3') which matched the consensus 5'-YAG-3' (Ballance, Transformation Systems for Filamentous Fungi and an Overview of Fungal Gene Structure, In Leong, S.A. and Berka, R.M. (eds.), *Molecular Industrial Mycology-Systems and Applications for Filamentous Fungi*, Dekker, Inc., New York, NY, pp. 1-29, 1991, 1986; Gurr et al., The Structure and Organization of Nuclear Genes of Filamentous Fungi, In Kinghorn (ed.), *Gene Structure in Eukaryotic Microbes*, IRL Press, Oxford, WA, pp.

93-139, 1987). The intron also contained an internal lariat sequence (5'-GCTGAC-3'; SEQ ID NO: 9), located 15 bp upstream of the 3'-splice site, consistent with the consensus 5'-(G/A)CT(G/A)AC-3'(SEQ ID NO: 19) for fungal introns.

The deduced amino acid sequence of the *A. nidulans* α-1,2-mannosidase 1A gene was aligned with other published α-mannosidases to determine whether there might have been other introns which did not shift the reading frame and did not contain stop codons but did increase the size of the putative gene product. Two additional regions of the first ORF did not align with other published sequences and appeared as large 'gaps' in the multiple sequence alignment for all of the other α-1,2-mannosidase genes used in the alignment. To determine whether these sequences represented introns or encoded polypeptide sequence novel to the *A. nidulans* α-mannosidase, RT-PCR products were compared to genomic PCR products. The PCR amplification products derived from both RNA and genomic DNA, using primer pairs which flanked these regions, were the same size, indicating that there were no introns present in these regions. Although introns can sometimes be present with less conserved splice sequences (Gurr et al., The Structure and Organization of Nuclear Genes of Filamentous Fungi, In Kinghorn (ed.), *Gene Structure in Eukaryotic Microbes*, IRL Press, Oxford, WA, pp. 93-139, 1987), these regions did not contain consensus splice motifs.

The α-1,2-mannosidase 1B gene contained three open reading frames separated by two regions which contained several stop codons and caused a shift in the reading frame. These regions were searched for consensus intron splice sites. Both of the putative introns contained consensus 5'- and 3'- splice sites and a consensus internal lariat sequence. The two putative introns were verified by RT-PCR. In both cases, the RNA-derived amplification product was smaller than the DNA-derived amplification product, and the size difference corresponded to the predicted size of each intron. Sequencing of the RT-PCR products verified the presence of a 51 bp and 53 bp intron at the respective splice sites. While the position of each intron in the coding region correlated directly to two of the introns found in the *P. citrinum* α-1,2-mannosidase gene, the intron sequences themselves were not conserved. The α-1,2-mannosidase 1C gene contained a single contiguous open

reading frame and did not contain any consensus intron sequences. This gene did not appear to contain any intron sequences.

The authentic start codon (ATG) of the α-1,2-mannosidase 1A gene was inferred by sequence context of the putative start codons, combined with protein sequence alignments with known α-mannosidase proteins. A potential start codon (ATG) occurred in frame 42 bp after the stop codon which defined the 5' end of the first ORF, while the next in frame ATG codon was 423 bp downstream. This first codon was thus a better candidate for the translational start codon. Translation originating at this start codon would produce a protein product with an N-terminus which was larger than other fungal α-mannosidases, but similar in size to the N-termini of mammalian and insect Class I α-mannosidases. Although there is not a strong consensus sequence surrounding the translational start codon in filamentous fungi (Ballance, *Yeast* **2**:229-236, 1986), there is a preference for a purine at position -3 (97%). The putative start codon for this gene had a G at position -3, and thus conforms to this rule. Placement of the translational start codon can also be inferred from sequence context with respect to promoter elements and the transcription start site. There was a TATA-like element at position -47 of the putative start site and several pyrimidine rich blocks within 100 bp upstream of the proposed start codon. These pyrimidine rich regions are often found in fungal promoters and may influence the level of transcription (Ballance, *Yeast* **2**:229-236, 1986). The 5' non-translated region did not contain a CAAT-box upstream of the TATA-like element, but this is not unusual for fungal promoters.

The first potential translational start codon in the α-1,2-mannosidase 1B gene occured 42 bp into the first ORF. Comparison of the coding region of this ORF with the *A. satoi* and *P. citrinum* α-1,2-mannosidase genes showed that the position of the putative translational start site correlated with the start sites of these genes. This start codon also contained a purine at the -3 position, a TATA-like element at position -76, and several CT-rich blocks in the sense strand. The first potential translational start codon of the α-mannosidase 1C gene occurred 38 bp into the ORF. The start site also contains a purine at position –3 and CCAAT motif at –221, but did not contain a clearly definable TATA box.

The putative coding region of the α-1,2-mannosidase 1A gene encodes a 816 amino acid protein with a predicted molecular weight (MW) of 91 kD. This is somewhat larger than other Class I α-mannosidases, which range in size from 53 kD for the *P. citrinum* α-mannosidase I (Yoshida et al., *Biochim. Biophys. Acta* **1263**:159-162, 1995) to 73 kD for the *H. sapiens* α-mannosidase 1B (Bause et al., *Eur. J. Biochem.* **217**:535-540, 1993), *M. musculus* α-mannosidase 1A (Lal et al., *J. Biol. Chem.* **269**:9872-9881, 1994) and α-mannosidase 1B (Herscovics et al., *J. Biol. Chem.* **269**:9864-9871, 1994), and *S. scrofa* α-mannosidase I (Bieberich et al., *Eur. J. Biochem.* **246**:681-689, 1997). The coding region of the α-1,2-mannosidase 1B gene encodes a 505 aa protein with a predicted MW of 56 kD, while the α-1,2-mannosidase 1C gene encodes a 586 aa protein with a predicted MW of 65 kD. Both of these predicted sizes are within the range of currently identified Class I α-mannosidases.

The putative α-1,2-mannosidase 1A, 1B, and 1C proteins contained several charged N-terminal amino acids representative of a typical signal sequence motif downstream of the putative start site. Kyte-Doolittle hydropathy plots showed that the signal sequences were followed by highly hydrophobic regions approximately 15-16 amino acids in length, while the remainder of the C-termini were relatively hydrophilic. These proteins likely form type II transmembrane proteins, which is a characteristic of other Class I α-1,2-mannosidases.

Example 7

Mannosidase activity assays: Crude protein extracts were obtained from protoplasts and from culture filtrates. Secreted proteins were precipitated from 1.5 mL of culture filtrate by saturation with ammonium sulfate. After centrifugation, the protein extract was resuspended in 500 mL 0.01 M phosphate buffer (pH 6.0) and precipitated again using ammonium sulfate (to ensure removal of sugar residues which would interfere with mannosidase assays). Proteins were resuspended in 100 μL 0.01 M phosphate buffer (pH6.0) and 27 μL of the extract was used in the mannosidase assay. Intracellular proteins were extracted from protoplasts, which were prepared as in Eades et al., *Glycobiol.* **8**:17-33, 1998. Protoplasts were

centrifuged and resuspended in 200 µL 0.01 M phosphate buffer (pH6.0)/1% octylthioglucoside, and resuspensions were vortexed vigorously to ensure complete protoplast lysis. Cellular debris was removed by centrifugation, and 27 µL of the lysate was used in mannosidase assays.

Mannosidase assays were performed using the disaccharide Man-α-1,2-Man-α-OCH_3 as a substrate in a coupled enzyme assay as described earlier (Scaman et al., *Glycobiol.* **6**:265-270, 1996), with some modifications. Digestion of the substrate was performed in a 30 µL final volume containing 27 µL of crude extract in 0.01M phosphate buffer (pH6.0) and 3 mL 100 mM disaccharide Man-α-1,2-Man-α-OCH_3 incubated at 37°C for 3 hours. Detection of released mannose was achieved by addition of 30 µL Tris-HCl (pH7.6) and 240 µL of develping solution, containing glucose oxidase (55 U/mL), horseradish peroxidase (1 U/mL) and o-dianisidine dihydrochloride (70 µg/mL), incubated at 37°C for 3 hours. Absorbance measurements at 450 nm determined final color change. Standard blanks included all components of the colrimetric reaction, plus the substrate. As a control, enzyme extracts which were not used in the mannosidase digestion were subjected to the colorimetric reaction, to determine the absorbance which is due to the extract itself, and not due to mannose release. These values were subtracted from the absorbance values of the assays. Free mannose was used as a standard. All assays were performed in triplicate and the mean and standard deviation was calculated for each sample. Mannosidase activity was standardized by comparison with total protein in the crude enzyme extracts, and was defined as the amount of mannose released from substrate per µg of total protein per hour. Protein concentrations were determined by the Bradford method (Bradford, *Anal. Biochem.* **72**:248-254, 1976) using BSA as a standard.

The method described above was used to verify that α-1,2-mannosidases (1A, 1B, and 1C; SEQ ID NOS: 3, 6, and 18, respectively) were actively expressed intracellularly. These assays were performed to determine the secreted and/or intracellular α-1,2-mannosidase activity levels. To determine secreted α-1,2-mannosidase activity, crude protein extracts were obtained from liquid culture filtrates. Intracellular proteins were extracted from fungal protoplasts using a

detergent buffer. Significant α-1,2-mannosidase activity was found in the intracellular protein extract (28.16 nmol mannose released from substrate/μg total protein/h; SD = 4.45), whereas very little activity was found in the extracellular extracts (0.77 nmol/μg/h; SD = 0.74). This is consistent with the hypothesis that the Class I α-1,2-mannosidase genes from *A. nidulans* encode type-II transmembrane proteins which are expressed intracellularly.

Example 8

Enzyme and protein assays. La et al., *Glycobiol.* **8:**981-995, 1998, described a protocol that can be used to assay for mannosidase activity. This protocol involved assays for testing the activity of two separate mannosidases.

Briefly, La et al., Glycobiol. 8:981-995, 1998, assayed for mannosidase activity using the disaccharide Man-α-1,2-manα-O-CH_3, as a substrate and the glucoseoxidase/peroxidase coupled enzyme assay as described earlier (Scaman et al., *Glyco. Biol.* **6:**265-270, 1996) with some modifications. Assays testing mannosidase activity in crude culture medium, or purified enzyme preparations were performed in flat-bottomed 96-well microtiter plates in a 25-μL total reaction volume containing 0.01 M potassium phosphate, pH 6.0, 2-10 mM Man-α-1,2-man-α-O-CH_3 and 5-15 mL of enzyme solution at 37°C for 30 minutes, or as specified. The reaction was terminated by adding 25 μL of 1.25 M Tris-CL, pH 7.6.

The La et al., *Glycobiol.* **8:**981-995, 1998, also provided an additional assay, that was used to test the activity of a second mannosidase. This second mannosidase was assayed in culture medium following 10-fold concentration using a Centricon 30 ultrafiltration membrane (Amicon). Assays were performed in 0.5-μL microfuge tubes in 30 μL total reaction volume containing 0.01 M potassium phosphate, pH 6.0, 2-10 mM Manα1,2 Manα-O-CH_3, and 15-24 μL of concentrated culture medium at 37°C for 2 hours, or as specified. The reaction was terminated by heating to 100°C for 5 minutes followed by centrifugation at 16,000 x *g* for 5 minutes. The supernatants (25 μL) were transferred to microtiter plate wells containing 25 μL of 1.25 M Tris-CL, pH 7.6. In both cases, the amount of mannose released was detected by incubation with 250 μL of "developing solution" containing glucose oxidase (55 U/mL), horseradish peroxidase (1 purpurogallin unit/mL), and *o*-

dianisidine dihydrochloride (70 μg/mL) for 3 hours and 37°C. The final color intensity was determined by measuring the absorbance at 450 nm on a Bio-Tek (Winooski, VT) microtiter plate reader. Free mannose was used as a standard. One unit of mannosidase activity is defined as the amount of enzyme that releases 1 nmol of mannose in 1 minute at 36°C.

Example 9

α1,2-Mannosidase activity assays. Herscovics, and Jelinek-Kelly, *Anal. Biochem.* **166:**85-89, 1987, describe an assay that can be used to test mannosidase activity from cell lysates. This assay can be used to test the activity of mannosidases of the present invention.

Briefly, lysates from Sf9 cells (Acc. # CRL-1711) were used for assays of α-1,2-mannosidase activity with [^{3}H]Man$_9$GlcNAc as described previously (Herscovics, and Jelinek-Kelly, *Anal. Biochem.* **166:**85-89, 1987) with some modifications. The cells were seeded into 25-cm^2 tissue culture flasks (Corning Glass Works, Corning, NY) at a density of 3 million cells per flask and infected at a multiplicity of 10 plaque forming units (PFU) per cell with either a recombinant baculovirus encoding the Sf9 α-1,2-mannosidase cDNA under the control of the polyhedrin promoter, or wild-type baculovirus as a control. The cells were harvested by centrifugation at selected times after infection, washed with 100 mM Na^+-MES (pH 6.0), and lysed using 1 μL per 30,000 cells of the same buffer containing 0.5% Triton X-100 with or without 80 mM EDTA. Five microliters of the lysates were used in each reaction of the α-1,2-mannosidase activity assays. Cell lysates for negative control reactions were carried out in a total volume of 40 μL containing 75 mM Na^+-MES (pH 6.0), 0.0625% Triton X-100, and 6000 c.p.m. of [^{3}H]Man$_9$GlcNAc. In addition, some reactions contained $CaCl_2$, $MgCl_2$, EDTA. The reactions were incubated at 37°C for 2.5 hours, boiled for 2 minutes, and 250 μL of a 5 mM solution of $CaCl_2$, $MgCl_2$, and $MnCl_2$, and 175 μL of a 2.25 μg/μL solution of concanavalin A (Boehringer-Mannheim, Indianapolis, IN) in 3.8 M NaCl were added to each reaction. The solutions were vortexed gently and incubated at room temperature for two minutes. Then, 1 mL of 25% PEG-8000 was added and the solutions were vortexed and incubated for another 5 minutes at room

temperature. The solutions were centrifuged for 2 minutes at 10,000 x *g*, 1 mL of the supernatant was added to 4 mL of a scintillation cocktail (Ultima Gold; Packard Instrument Company, Meriden, CT), and radioactivity was measured using a liquid scintillation counter (model LS 6000Icl Beckman Instruments, Inc., Fullerton, CA). (Herscovics and Jelinek-Kelly, *Anal. Biochem.* **166**:85-89, 1987).

Strains, Media and Growth Conditions: *Aspergillus nidulans* sporecolor mutant SM222 was grown in CYM liquid medium (10 g glucose; 2 g bactopeptone; 1.5 g casamino acids; 1 g yeast extract; 10 mL 100X salt solution, 1 mL 1000X trace elements, 10 mL 100X vitamin solution, and 10 mL 100X adenine solution per litre). Stock solutions (100X salt, 100X vitamin, 1000X trace elements, 100X adenine) were described in Kalsner et al., *Glycoconjugate J.* **12**:360-370, 1995. Strains were maintained on CYM agar, and spore suspensions were obtained by washing cultured CYM agar plates with 8 mL 0.001% Tween 80. Mycelia for DNA isolations were obtained by inoculating 500 mL liquid CYM agar with 10^8 spores, and incubating 24 hours at 30°C with constant agitation (200 rpm).

Oligonucleotide Primer Design: The forward primer MANFOR2B was designed by reverse translation of the protein sequence GGLGESFYEY (SEQ ID NO: 13), and the reverse primer MANREV3B was designed from the complement of the reverse translation of the sequence FXLAETLKYLY (SEQ ID NO: 14). These protein sequences were conserved between the α-1,2-mannosidase protein sequences of *S. cerevisiae* (Camirand et al., *J. Biol. Chem.*, **266**:15120-15127, 1991), *Mus musculus* 1A (Lal et al., *J. Biol. Chem.*, **269**:9872-9881, 1994; Acc. # U04299), *Homo sapiens* 1A (Bause et al., *Eur. J. Biochem.*, **217**:535-540, 1993; Acc. # X74837), and rabbit liver (Lal et al., *J. Biol. Chem.*, **269**:9872-9881, 1994). A codon usage table compiled for *A. nidulans* was used to aid in nucleotide selection at degenerate sites.

DNA Isolation and PCR Amplification: Total genomic DNA was extracted from finely ground freeze-dried mycelia of *A. nidulans* strain SM222. Approximately 400 mg of mycelia were vortexed with 2.5 mL of 50 mM EDTA,

0.2% SDS, and centrifuged for 10 minutes. Then, 85 μL of 3 M KOAc, 5 M acetic acid were added to the supernatant. Following a 20-minute incubation on ice, the suspension was re-centrifuged and DNA was isopropanol-precipitated from the supernatant. After resuspension in 100 μL TE (10 mM Tris pH 7.5, 1 mM EDTA), the DNA was extracted once with phenol, twice with chloroform/isoamyl alcohol (24:1) and ethanol precipitated. Each PCR reaction consisted of 10-100 ng of genomic DNA, 50 pmol of each primer, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl, 50 mM KCl, 0.01% gelatin, 0.1% Triton X-100, 200 μM each of dATP, dCTP, dTTP, and dGTP, and 2 units *Taq* DNA polymerase (Perkin-Elmer, Norwalk, Connecticut) in a final volume of 100 μL. Amplification was performed in two stages using a Perkin-Elmer thermal cycler. Five cycles at a lower stringency (56°C) were followed by 30 cycles at higher stringency (63°C). The PCR products were eluted from 1% low-melting-point agarose, cloned into T-Vector (Promega, Madison, Wisconsin) using the T-Vector cloning system and sequenced with Universal Forward and M13 Reverse primers (Amersham International, Buckinghamshire, England).

Library Construction and Screening: The cloned PCR products were used as a template to produce a radiolabeled PCR probe for library screening. Briefly, the PCR reactions were performed as previously described, except that only 100 μM dCTP was used and 0.825 uL α-^{32}P-dCTP (5 μCi) was added to the reactions (reduction of the dCTP concentration prior to addition of the radiolabeled nuclide reduces error due to base misincorporation). The radiolabeled PCR products were purified using the Wizard™ PCR Prep Kit (Promega) and scintillation counted to assess radioactivity.

A genomic library of *A. nidulans* (SM222) sequences was constructed by digesting genomic DNA with *Bam*HI and ligating the resulting fragments into the similarly digested lambda DNA vector EMPL-3. Concatemers of the ligated DNA were packaged using the Gigapack™ II (Stratagene, La Jolla, California) *in vitro* packaging system. Approximately 10^5 recombinant lambda plaques were immobilised on nylon membranes (Genescreen Plus™, Dupont, Wilmington, Delaware) and hybridized with the radiolabeled PCR product generated from *A.*

nidulans. Single hybridizing clones were selected and rescreened. Lambda DNA was isolated with the Wizard™ Lambda DNA Extraction Kit (Promega), digested with restriction enzymes, and subcloned into the pUC18 cloning vector.

DNA Preparation and Sequencing of *A. nidulans* α-1, 2-mannosidase 1A: Subclones, from the *A. nidulans* library, that contained the full-length α-1,2-mannosidase genes were identified by Southern analysis and sequenced. Initial manual sequencing of the α-mannosidase 1A gene was performed according to the dideoxynucleotide method using the T7 sequencing kit (Pharmacia). Initial sequence data were obtained with the universal priming sites, and with specific sequence primers (primer walking). Final sequence data were provided by subcloning the fragments using various restriction enzymes and sequencing with an ABI373 automated fluorescent sequencer (Applied Biosystems, Foster, California). Open reading frames were identified and aligned with known α-1,2-mannosidase sequences. The α-mannosidase 1B gene was sequenced by cloning restriction fragments of the positive lambda subclone into BlueScript II™ cloning vector (Stratagene, La Jolla California) and sequencing on an ABI373 sequencer, using the universal priming sites of the vector.

Identification of Introns: Total RNA was extracted from fresh mycelia grown in 150 mL liquid media using the guanidine isothiocyanate method (TRIzol, Molecular Research, Cincinnati, Ohio). Approximately 1 gram of fresh mycelia harvested by suction filtration was homogenized with 6 mL of TRIzol reagent, and incubated at 30°C for 10 minutes. To the homegenate, 1.2 mL chloroform was added, and the mixture was shaken vigorously for 15 seconds, incubated at 30°C for 2 minutes and centrifuged at 3000 rpm. The upper aqueous phase was removed and 3 mL isopropanol were added to precipitate the RNA. After incubation at 30°C for 10 minutes, the RNA was pelleted by centrifugation and washed with 70% ethanol/diethylpyrocarbonate (DEPC). The RNA pellet was resuspended in 30 μL H_2O/DEPC and 5 μL were used for cDNA preparation and PCR amplification.

Reverse-transcriptase PCR was performed using Superscript II™ reverse transcriptase (Gibco BRL) for first-strand cDNA synthesis using an oligo-dT primer

followed by PCR amplification of the cDNA. The oligo-dT primer was annealed to the RNA by adding 1 μL of oligo-dT primer (500 μg/mL) to 5 μL of RNA (approx. 1-3 μg) and 6 μL of H_2O/DEPC. The mixture was heated to 70°C for 10 minutes and then quickly chilled on ice. Reverse transcription was achieved by addition of 4 μL "First Strand Buffer" (250 mM Tris-HCl pH 8.3, 375 mM KCl, 15 mM $MgCl_2$), 2 μL 0.1 M dithiothrectol (DTT), 1 μL 10 mM dNTPs, and 1 μL (200 U) Superscript II™ reverse transcriptase, followed by incubation at 42°C for 50 minutes. After first-strand cDNA synthesis, the RNA was digested by the addition of 1 μL (2 U) RNAase H, and incubated at 37°C for 20 minutes. The cDNA was used to amplify regions containing putative introns for comparison with genomic DNA amplification.

Sequence Comparison: The Class-1 α-mannosidase protein sequences used in the sequence comparisons were obtained from GenBank as follows: *Saccharomyces cerevisiae* (Camirand et al., *J. Biol. Chem.*, **266**:15120-15127, 1991; Acc. # M63598), *Aspergillus satoi* (Inoue et al., *Biochim. Biophys. Acta,* **1253**:141-145, 1995; Acc. # D49827), *Penicillium citrinum* (Yoshida and Ichishima, *Biochim. Biophys. Acta,* **1263**:159-162, 1995; Acc. # D45839), *Ophiostoma novo-ulmi* (C.J. Eades and W.E. Hintz, unpublished; Acc. # AF129495), *Drosophila melanogaster* (Kerscher et al., *Dev. Biol.*, **168**:613-626, 1995; Acc. # X82640), *Spodoptera frugiperda* (Kawar et al., *Glycobiol.* 7:433-443, 1997; Acc. # AF005035*). Homo sapiens* 1A (Bause et al., *Eur. J. Biochem.*, **217**:535-540, 1993; Acc. # X74837), *Homo sapiens* 1B (Tremblay et al., *Glycobiol.* **8**:585-595, 1998; Acc. # AF027156), *Mus musculus* 1A (Lal et al., *J. Biol. Chem.* **269**:9872-9881, 1994; Acc. # U04299), *Mus musculus* 1B (Herscovics et al., *J. Biol. Chem.*, **269**:9864-9871, 1994; Acc. # U03457), and *Sus scrofa* (Bieberich et al., *Eur. J. Biochem.*, **246**:681-689, 1997; Acc. # Y12503). These sequences were aligned using the ClustalW algorithm (Thompson et al., *Nuc. Acid. R.* **22**:4673-4680, 1994) included in the DNAStar computer package (DNAStar, Madison Wisconsin). The sequence similarity matrix and dendrogram were also generated with DNAStar.

Having illustrated and described the principles of the invention in multiple embodiments and examples, it should be apparent to those skilled in the art that the

invention can be modified in arrangement and detail without departing from such principles. We claim all modifications coming within the spirit and scope of the following claims.

We Claim:

1. A purified protein having mannosidase activity, comprising an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence selected from the group consisting of SEQ ID NOS: 3, 6 and 18;

(b) an amino acid sequence that differs from those specified in (a) by one or more conservative amino acid substitutions; and

(c) amino acid sequences having at least 60% sequence identity to the sequences specified in (a).

2. A specific binding agent, that binds to the purified protein of claim 1.

3. An isolated nucleic acid, molecule encoding a protein according to claim 1.

4. A recombinant nucleic acid molecule, comprising a promoter sequence operably linked to a nucleic acid sequence according to claim 3.

5. A cell, transformed with a recombinant nucleic acid molecule according to claim 4.

6. The transformed cell of claim 6, wherein the cell is selected from the group consisting of: an insect cell, a yeast cell, an algae cell, a bacterial cell, a mammalian cell, and a plant cell.

7. A transgenic fungus, comprising a recombinant nucleic acid according to claim 4.

8. A method for altering the glycosylation pattern of a macromolecule, comprising contacting the macromolecule with a purified protein according to claim 1.

9. The method of claim 8, wherein contacting the macromolecule is performed *in vitro*.

10. A method for producing a macromolecule having an altered glycosylation pattern, comprising:

(a) providing a transformed cell according to claim 4; and

(b) allowing the transformed cell to produce the macromolecule.

11. An isolated nucleic acid molecule, comprising a sequence selected from the group consisting of:

(a) at least 15 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;

(b) at least 20 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;

(c) at least 30 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;

(d) at least 15 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;

(e) at least 20 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;

(f) at least 30 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;

(g) at least 15 contiguous nucleotides of the sequence shown in SEQ ID NO: 17;

(h) at least 20 contiguous nucleotides of the sequence shown in SEQ ID NO: 17; and

(i) at least 30 contiguous nucleotides of the sequence shown in SEQ ID NO: 17.

12. A method for altering the glycosylation pattern of a macromolecule in a sample, comprising:

(a) adding a purified protein according to claim 1 to the sample;

(b) incubating the sample with the purified protein; and

(c) allowing the purified protein to hydrolytically remove at least one terminal mannoside residue from a macromolecule in the sample.

13. A method for isolating a nucleic acid sequence encoding a mannosidase, comprising:

(a) hybridizing the nucleic acid sequence under high-stringency conditions to at least 50 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 1, 4, and 17; and

(b) identifying the nucleic acid sequence as one that encodes a mannosidase.

14. A mannosidase identified by the method of claim 13.

Abstract

Mannosidase enzymes and use of such enzymes to alter the glycosylation patterns of macromolecules are disclosed. Also disclosed are the nucleic acid sequences encoding the mannosidase enzymes.

SLR:kdn 03/25/02 2847-62447 106857 doc

PATENT
Attorney's Matter No. 2847-62447

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: William E. Hintz and Caleb Joshua Eades
Application No. Not yet assigned
Filed: herewith
For: MANNOSIDASES AND METHODS FOR USING SAME

Art Unit: Not yet assigned

Examiner: Not yet assigned

Date: March 25, 2002

STATEMENT IN COMPLIANCE WITH 37 C.F.R. § 1.821(f)

COMMISSIONER FOR PATENTS
Washington, DC 20231

In compliance with 37 C.F.R. § 1.821(f), the undersigned declares that the nucleotide and/or amino acid sequences presented in the paper copy of the "Sequence Listing" submitted herewith are the same as the sequences contained in the computer-readable form of said "Sequence Listing." No new matter has been added.

Respectfully submitted,

KLARQUIST SPARKMAN, LLP

By _____
Sheree Lynn Rybak, Ph.D.
Registration No. 47,913

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391
Facsimile: (503) 228-9446

SEQUENCE LISTING

<110> Hintz et al.

<120> Mannosidases and Methods for using the Same

<130> 62447

<140>
<141>

<150> PCT/US00/27210
<151> 2000-10-02

<150> 60/157,341
<151> 1999-10-01

<160> 19

<170> PatentIn version 3.1

<210> 1
<211> 3328
<212> DNA
<213> Aspergillus nidulans

<400> 1
agggcctttt caggggctgg ttggccaaac ctgccgtatt ttcgattcag ggcgggccct 60

taattctggg gaccgatatc ccttgaacgc gggggcaatc agttcaacca acccagaccg 120

ctgggcttgg tatgagattg cgagcatcta tgtcggtgta ccttctgaat gacaatgaat 180

gtattttact tctcgaaaag aacccttggg cactgaattg tgcggagaat gatgccctga 240

ttatgataca actagtccgc tccgtcaagc cacaagggtc tgggcagtcc gctataaatc 300

aaaatcgcct gcacgaacag acgaataacc aagaaaacgc ccgagcgcga gcgtttcttc 360

ttcctctaag ccttgcagct ggctctgcgt ctttgatcaa ccctttagct gaatttcccc 420

agaacttcag ccctctgcat cctgtcctta ccgcaactcg ttaacctgcg cgacctcgcg 480

cgaccacagc cttaggtttc gagatgccat gaaaatcaga aattgaaccc cctttccatt 540

actatcattc tctgcattct gcgagtgatc tgtccttcga cgttccttct ttccagcgct 600

gcggcgcctt cactctcgtt gcctacgttt gaccacggtc ctacctctcc tactgctgat 660

tattaggctc ctccctacgc ctccaataca gggaagtcgc cggccatgtt tcgtgcacga 720

cgatctcgca tctcgctggt gtttgccgtt atatttgtcc tcctcatatt ccactttagc 780

cgtctcgcag ttacgatcag cctgcaatct tgggtacctc cgccgcccgt cgatcaccat 840

aatccccctt tccccgacca gaacctcaaa gatccatacg aaaacgacaa tagtgcgacc 900

ggcagtgggg ctcctccgcc tgcgttggta gagccagaag aataccaacg accaccactt 960

tacacagatt cagatgacag cccaactccg tcaaaagaac gcctggacac cccgagcaat 1020

gtcccatctc aggagcctga atttgatgcc gccagacttc agacgggtgc gcagacccaa 1080

aataaacatg aagatgatga ggatattgtc ccaatttctc actggaagcc gatgcccgaa 1140

cggcatccag tcagtccgga ggctttgatc aagctgccaa ccgggcaatc aaaggaactc 1200

ccccaactgc aagctaagtt caaggacgag tcgtcctcgg acaagatgca gcggctgcaa 1260

caacttgaca ctatcaagtc ggcgttctta catgcgtgga acggttacaa gatctctgcc 1320

atgggtcatg atgaggttag acctctgcgc ggtggtttca aggacacatt caatggctgg 1380

ggcgcgaccc ttgtcgacgc cttggatacc ctgtggatca tggatctcaa agaggagttc 1440

tccatggcag tcgactacgt caagaaaatc gattttacca ccagcaccaa gaaagagatt 1500

ccggtctttg aaaccactat tcgctaccta ggcgggatgc tcggggccta tgatatttcg 1560

ggacacaaat acgatatact tttggaaaag tctgttgagc ttgcggatgt cttgatggac 1620

gccttcgaca caccgaaccg gatgccaacc ctctattata aatggagccc agagtatgct 1680

tcagagtttc gccgggggga ctttaaggct gttctcgccg agcttggctc tctctctctc 1740

gagttcacgc gtttggcgca gttgaccaaa caggacaagt actacgatgc aattgcacga 1800

atcacaaatg agctcgaaaa gtatcaggat ttgacaaagc ttcccggctt gtggcctctc 1860

aacctggacg catccgggtg caggcgagtt cccggcgtct cgcgagagcc tgctgcggct 1920

gggcagccag tcagatggtc ctctgacgag atcaactcga cgagctcggt atcgtatcgt 1980

acaagacaaa ttcatgaggg cggagagcct gtccgtcatg acaatgattc gtttgaaacg 2040

ggttttcctg tatcagtcga tactcggact cctcccccaa agcaagattg caccggaggc 2100

ctcaacgatc agctctcagg cattgacaag ttcggactcg gagcccttgg tgactctacg 2160

tacgagtact taccgaaaga gtatatgttg ctcggcggta acaacgacca gtacctcaac 2220

atgtatcaga aggccatgga cacagtgcga gaatatcttg tttatcagcc aatgctcaag 2280

aataatcgcg atgtccgctt cttagcgaca gttagtatga caaagagcct tgatgcaaac 2340

cctccggggc gtaccacttt cgcgtacgaa ggcactcacc tcacctgttt tgctggtggt 2400

atgcttgcca ttggcgccaa gttgtttggg cttgataagg atctaaagct gggtagtcaa 2460

ctgacggacg gctgtgtctg ggcatatgaa gccacaaagt ccggaatcat gccggaagca 2520

ttccaactgg tcccttgtaa gaaaggcgag ccatgcgaat gggatgagga cgcatactac 2580

atggccatgg atccttatgc cgacaagcgg ccaatatcac ataacaaacg ctccgccggc 2640

cctgaaaagg ggaattggca cgtcgtcgcc acagccgaat cgtcttcgcc ccaggaagat 2700

aaaacacaga aatcaaccac tactgagggt cgacacaccg gtacaactac cggggcaggc 2760

gcgctctcgc acgaggaatt cgtcacggga aaaatcctca acgaccgact cccgccgggc 2820

atgacaggga tctcggctcg gcagtacctc cttcgcccgg aggcgatcga gtctgtcttc 2880

atcatgttcc gcctcacggg cgatccttcc tggcgcgaaa agggttggaa gatgttccag 2940

gctgtcgaca aagccacgaa gacggagctg gcgaactcgg ccatttccga cgtaaccgtc 3000

gataatccac gcccggtgga cagtatggaa tcattctggc ttgcggagac tctgaaatac 3060

ttctaccttc ttttcagcga tccaagcctg gtgagccttg acgaatatgt cttgtaagtg 3120

atgcttgact taatcgactg cttgatgctg acttttccct taggaacacc gaggctcatc 3180

cgttcaagcg acccaagtac tgaagtacta atttaaatga tcttttagcc tgtatctata 3240

catggccgct ccgctgtaga agcattgata ccattaagac agtatcgctg cattcgtgta 3300

ccatttgagc ttccagagga acctcttt 3328

<210> 2
<211> 2448
<212> DNA
<213> Aspergillus nidulans

<220>
<221> variation
<222> (1632)..(1632)
<223> R = A, C, G, or T

<400> 2
atgtttcgtg cacgacgatc tcgcatctcg ctggtgtttg ccgttatatt tgtcctcctc 60

atattccact ttagccgtct cgcagttacg atcagcctgc aatcttgggt acctccgccg 120

cccgtcgatc accataatcc ccctttcccc gaccagaacc tcaaagatcc atacgaaaac 180

gacaatagtg cgaccggcag tggggctcct ccgcctgcgt tggtagagcc agaagaatac 240

caacgaccac cactttacac agattcagat gacagcccaa ctccgtcaaa agaacgcctg 300

gacaccccga gcaatgtccc atctcaggag cctgaatttg atgccgccag acttcagacg 360

ggtgcgcaga cccaaaataa acatgaagat gatgaggata ttgtcccaat ttctcactgg 420

aagccgatgc ccgaacggca tccagtcagt ccggaggctt tgatcaagct gccaaccggg 480

caatcaaagg aactccccca actgcaagct aagttcaagg acgagtcgtc ctcggacaag 540

atgcagcggc tgcaacaact tgacactatc aagtcggcgt tcttacatgc gtggaacggt 600

```
tacaagatct ctgccatggg tcatgatgag gttagacctc tgcgcggtgg tttcaaggac    660
acattcaatg gctggggcgc gacccttgtc gacgccttgg ataccctgtg gatcatggat    720
ctcaaagagg agttctccat ggcagtcgac tacgtcaaga aaatcgattt taccaccagc    780
accaagaaag agattccggt ctttgaaacc actattcgct acctaggcgg gatgctcggg    840
gcctatgata tttcgggaca caaatacgat atacttttgg aaaagtctgt tgagcttgcg    900
gatgtcttga tggacgcctt cgacacaccg aaccggatgc caaccctcta ttataaatgg    960
agcccagagt atgcttcaga gtttcgccgg ggggacttta aggctgttct cgccgagctt   1020
ggctctctct ctctcgagtt cacgcgtttg gcgcagttga ccaaacagga caagtactac   1080
gatgcaattg cacgaatcac aaatgagctc gaaaagtatc aggatttgac aaagcttccc   1140
ggcttgtggc ctctcaacct ggacgcatcc gggtgcaggc gagttcccgg cgtctcgcga   1200
gagcctgctg cggctgggca gccagtcaga tggtcctctg acgagatcaa ctcgacgagc   1260
tcggtatcgt atcgtacaag acaaattcat gagggcggag agcctgtccg tcatgacaat   1320
gattcgtttg aaacgggttt tcctgtatca gtcgatactc ggactcctcc cccaaagcaa   1380
gattgcaccg gaggcctcaa cgatcagctc tcaggcattg acaagttcgg actcggagcc   1440
cttggtgact ctacgtacga gtacttaccg aaagagtata tgttgctcgg cggtaacaac   1500
gaccagtacc tcaacatgta tcagaaggcc atggacacag tgcgagaata tcttgtttat   1560
cagccaatgc tcaagaataa tcgcgatgtc cgcttcttag cgacagttag tatgacaaag   1620
agccttgatg cnaaacctcc ggggcgtacc actttcgcgt acgaaggcac tcacctcacc   1680
tgttttgctg gtggtatgct tgccattggc gccaagttgt ttgggcttga taaggatcta   1740
aagctgggta gtcaactgac ggacggctgt gtctgggcat atgaagccac aaagtccgga   1800
atcatgccgg aagcattcca actggtccct tgtaagaaag gcgagccatg cgaatgggat   1860
gaggacgcat actacatggc catggatcct tatgccgaca agcggccaat atcacataac   1920
aaacgctccg ccggccctga aaaggggaat tggcacgtcg tcgccacagc cgaatcgtct   1980
tcgccccagg aagataaaac acagaaatca accactactg agggtcgaca caccggtaca   2040
actaccgggg caggcgcgct ctcgcacgag gaattcgtca cgggaaaaat cctcaacgac   2100
cgactcccgc cgggcatgac agggatctcg gctcggcagt acctccttcg cccggaggcg   2160
atcgagtctg tcttcatcat gttccgcctc acgggcgatc cttcctggcg cgaaaagggt   2220
tggaagatgt tccaggctgt cgacaaagcc acgaagacgg agctggcgaa ctcggccatt   2280
```

tccgacgtaa ccgtcgataa tccacgcccg gtggacagta tggaatcatt ctggcttgcg 2340

gagactctga aatacttcta ccttcttttc agcgatccaa gcctggtgag ccttgaggaa 2400

tatgtcttga acaccgaggc tcatccgttc aagcgaccca ggtactga 2448

<210> 3
<211> 815
<212> PRT
<213> Aspergillus nidulans

<400> 3

Met Phe Arg Ala Arg Arg Ser Arg Ile Ser Leu Val Phe Ala Val Ile
1 5 10 15

Phe Val Leu Leu Ile Phe His Phe Ser Arg Leu Ala Val Thr Ile Ser
20 25 30

Leu Gln Ser Trp Val Pro Pro Pro Pro Val Asp His His Asn Pro Pro
35 40 45

Phe Pro Asp Gln Asn Leu Lys Asp Pro Thr Glu Asn Asp Asn Ser Ala
50 55 60

Thr Gly Ser Gly Ala Pro Pro Pro Ala Leu Val Glu Pro Glu Glu Thr
65 70 75 80

Gln Arg Pro Pro Leu Thr Thr Asp Ser Asp Asp Ser Pro Thr Pro Ser
85 90 95

Lys Glu Arg Leu Asp Thr Pro Ser Asn Val Pro Ser Gln Glu Pro Glu
100 105 110

Phe Asp Ala Ala Arg Leu Gln Thr Gly Ala Gln Thr Gln Asn Lys His
115 120 125

Glu Asp Asp Glu Asp Ile Val Pro Ile Ser His Trp Lys Pro Met Pro
130 135 140

Glu Arg His Pro Val Ser Pro Glu Ala Leu Ile Lys Leu Pro Thr Gly
145 150 155 160

Gln Ser Lys Glu Leu Pro Gln Leu Gln Ala Lys Phe Lys Asp Glu Ser
165 170 175

```
Ser Ser Asp Lys Met Gln Arg Leu Gln Gln Leu Asp Thr Ile Lys Ser
            180                 185                 190

Ala Phe Leu His Ala Trp Asn Gly Thr Lys Ile Ser Ala Met Gly His
        195                 200                 205

Asp Glu Val Arg Pro Leu Arg Gly Gly Phe Lys Asp Thr Phe Asn Gly
    210                 215                 220

Trp Gly Ala Thr Leu Val Asp Ala Leu Asp Thr Leu Trp Ile Met Asp
225                 230                 235                 240

Leu Lys Glu Glu Phe Ser Met Ala Val Asp Thr Val Lys Lys Ile Asp
                245                 250                 255

Phe Thr Thr Ser Thr Lys Lys Glu Ile Pro Val Phe Glu Thr Thr Ile
            260                 265                 270

Arg Thr Leu Gly Gly Met Leu Gly Ala Thr Asp Ile Ser Gly His Lys
        275                 280                 285

Thr Asp Ile Leu Leu Glu Lys Ser Val Glu Leu Ala Asp Val Leu Met
    290                 295                 300

Asp Ala Phe Asp Thr Pro Asn Arg Met Pro Thr Leu Thr Thr Lys Trp
305                 310                 315                 320

Ser Pro Glu Thr Ala Ser Glu Phe Arg Arg Gly Asp Phe Lys Ala Val
                325                 330                 335

Leu Ala Glu Leu Gly Ser Leu Ser Leu Glu Phe Thr Arg Leu Ala Gln
            340                 345                 350

Leu Thr Lys Gln Asp Lys Thr Thr Asp Ala Ile Ala Arg Ile Thr Asn
        355                 360                 365

Glu Leu Glu Lys Thr Gln Asp Leu Thr Lys Leu Pro Gly Leu Trp Pro
    370                 375                 380

Leu Asn Leu Asp Ala Ser Gly Cys Arg Arg Val Pro Gly Val Ser Arg
385                 390                 395                 400
```

Glu Pro Ala Ala Ala Gly Gln Pro Val Arg Trp Ser Ser Asp Glu Ile
405 410 415

Asn Ser Thr Ser Ser Val Ser Thr Arg Thr Arg Gln Ile His Glu Gly
420 425 430

Gly Glu Pro Val Arg His Asp Asn Asp Ser Phe Glu Thr Gly Phe Pro
435 440 445

Val Ser Val Asp Thr Arg Thr Pro Pro Pro Lys Gln Asp Cys Thr Gly
450 455 460

Gly Leu Asn Asp Gln Leu Ser Gly Ile Asp Lys Phe Gly Leu Gly Ala
465 470 475 480

Leu Gly Asp Ser Thr Thr Glu Thr Leu Pro Lys Glu Thr Met Leu Leu
485 490 495

Gly Gly Asn Asn Asp Gln Thr Leu Asn Met Thr Gln Lys Ala Met Asp
500 505 510

Thr Val Arg Glu Thr Leu Val Thr Gln Pro Met Leu Lys Asn Asn Arg
515 520 525

Asp Val Arg Phe Leu Ala Thr Val Ser Met Thr Lys Ser Leu Asp Ala
530 535 540

Asn Pro Pro Gly Arg Thr Thr Phe Ala Thr Glu Gly Thr His Leu Thr
545 550 555 560

Cys Phe Ala Gly Gly Met Leu Ala Ile Gly Ala Lys Leu Phe Gly Leu
565 570 575

Asp Lys Asp Leu Lys Leu Gly Ser Gln Leu Thr Asp Gly Cys Val Trp
580 585 590

Ala Thr Glu Ala Thr Lys Ser Gly Ile Met Pro Glu Ala Phe Gln Leu
595 600 605

Val Pro Cys Lys Lys Gly Glu Pro Cys Glu Trp Asp Glu Asp Ala Thr
610 615 620

Thr Met Ala Met Asp Pro Thr Ala Asp Lys Arg Pro Ile Ser His Asn

625 630 635 640

Lys Arg Ser Ala Gly Pro Glu Lys Gly Asn Trp His Val Val Ala Thr
645 650 655

Ala Glu Ser Ser Ser Pro Gln Glu Asp Lys Thr Gln Lys Ser Thr Thr
660 665 670

Thr Glu Gly Arg His Thr Gly Thr Thr Thr Gly Ala Gly Ala Leu Ser
675 680 685

His Glu Glu Phe Val Thr Gly Lys Ile Leu Asn Asp Arg Leu Pro Pro
690 695 700

Gly Met Thr Gly Ile Ser Ala Arg Gln Thr Leu Leu Arg Pro Glu Ala
705 710 715 720

Ile Glu Ser Val Phe Ile Met Phe Arg Leu Thr Gly Asp Pro Ser Trp
725 730 735

Arg Glu Lys Gly Trp Lys Met Phe Gln Ala Val Asp Lys Ala Thr Lys
740 745 750

Thr Glu Leu Ala Asn Ser Ala Ile Ser Asp Val Thr Val Asp Asn Pro
755 760 765

Arg Pro Val Asp Ser Met Glu Ser Phe Trp Leu Ala Glu Thr Leu Lys
770 775 780

Thr Phe Thr Leu Leu Phe Ser Asp Pro Ser Leu Val Ser Leu Glu Glu
785 790 795 800

Thr Val Leu Asn Thr Glu Ala His Pro Phe Lys Arg Pro Arg Thr
805 810 815

<210> 4
<211> 2177
<212> DNA
<213> Aspergillus nidulans

<400> 4
cggaatgtgc ctaaagtgga aggtatgatg atgcccagga tcgcgcccca gtcatcaact 60

ccatcatggg acggtccttg atcctcaagg cacgaagtgg agatcaggtc cgtagtgcat 120

atgcatggcc catcagcctg aagcacttcc ccaagcaaag tcgagactcg gacaccgatg 180

atatccctgc tgtccctgac tgatgcatag tgcatgcccc tgcgctggct ccccttttca 240

ctccgcctgg tctccagtct ccactcctca ccattgatgt ctgcccccgc ccgccctcca 300

tcctccatca ttcttatatc tacggactcg gtcactcgtt atcactagag tccttgttta 360

ttcctagtgt ttgcattctt acgtgtagtt atgcgtacgc ttctcgctct cgcggccttc 420

gcgggctttg ccgctgctag ggtgcccgcc tacgccatca cgcgcccggt gatgcgcagt 480

gattctcgcg ccgacgctgt caaggaggcc ttttcgcatg cctgggacgg ttactacaac 540

tacgcttttc ctcatgacga gcttcacccg atttcgaacg gttacggaga ctcgcgaaac 600

cactggggcg cgtcggccgt cgacgctcta tcgacggcca tcatgatgcg caacgcgacc 660

atcgtcaacc agatccttga ccatattgct gctgtggact actccaagac caacgccatg 720

gtaagtttgt tcgagacgac catccggtac ctcgcgggca tgatttctgg atacgacctg 780

ctcaaaggcc ctgcggcggg gttggtggac gacagcaggg tcgacgtgct tctagagcag 840

tcgcagaacc tcgccgaggt gctgaaattc gcgttcgaca ctccttctgg tgtgccgtac 900

aacatgatta acattacttc gggcggcaac gacggcgcca cgaccaacgg gctggccgtg 960

accggtacct tggtgctgga gtggacgcgt ctgtcggact tgactgggaa cgacgagtat 1020

gcccgcttga gccagagagc tgaagactac cttctccacc cggagccagc gcagtacgaa 1080

ccgttccctg gattgattgg aagcgcagtc aatattgccg acggcaagct cgccaatggt 1140

cacatcagct ggaatggtgg cgcagactcg tactacgagt acctgatcaa gatgtacgtc 1200

tacgatcccg aacgctttgg cctctaccgg gaccgctggg tcgcagctgc cgagtcgagc 1260

atcaaccatc tggcttcgca cccgtccacc cgcccagacg tgactttctt ggccacttac 1320

aacgaggagc atcagctggg cctgaccagc caacacctga cctgcttcga cggtggaagc 1380

tttctgcttg gtgggacatt gctggaccgc caggactttg tcgacttcgg ccttgacctt 1440

gtcgccggct gccacgagac gtacaactcg actctgacgg gcatcggccc tgagcaattc 1500

agctgggacc ctaacggtgt gcccgacagc cagaaggagc tgttcgagcg cgcaggcttc 1560

tacatcaaca gcggccaata cattcttcgt cccgaagtca tcgagagctt ctactatgca 1620

tggcgcgtca caggtgatgg aacggtacgt tcactcagcg ctgcttccgt aggaagacca 1680

tactgaccgc tttagtacct cgaatgggtg tggaacgcct tcaccaacat caacaagtac 1740

tgccgcactg cgaccggttt cgcggggctg gagaacgtca atgcagcgaa cggcggaggc 1800

cggatcgaca accaggagag tttcatgttc gcagaggtgc tgaagtattc gtttttgact 1860

tttgctcctg gtacgtttcc ctagttctgg ttcacctgtg gagaatatta ctgactgcag 1920

cagaggacga ctggcaggtg cagaagggca gtggaaatac gtttgtttat aacaccgagg 1980

cgcacccgtt taaggtgtat acgcctcagt agatagtaca tattgtgctc tagcgtactg 2040

catgcataac ctaaatggca tcacctactt actgactact ctactgacaa gcagttgcct 2100

ttatatgctc ttaaggagct tgggcggtaa atcgtaattc actctaaact ccgcgtgctg 2160

ttcggggaag gctgggc 2177

<210> 5
<211> 1515
<212> DNA
<213> Aspergillus nidulans

<400> 5
atgcgtacgc ttctcgctct cgcggccttc gcgggctttg ccgctgctag ggtgcccgcc 60

tacgccatca cgcgcccggt gatgcgcagt gattctcgcg ccgacgctgt caaggaggcc 120

ttttcgcatg cctgggacgg ttactacaac tacgcttttc ctcatgacga gcttcacccg 180

atttcgaacg gttacggaga ctcgcgaaac cactggggcg cgtcggccgt cgacgctcta 240

tcgacggcca tcatgatgcg caacgcgacc atcgtcaacc agatccttga ccatattgct 300

gctgtggact actccaagac caacgccatg gtaagtttgt tcgagacgac catccggtac 360

ctcgcgggca tgatttctgg atacgacctg ctcaaaggcc ctgcggcggg gttggtggac 420

gacagcaggg tcgacgtgct tctagagcag tcgcagaacc tcgccgaggt gctgaaattc 480

gcgttcgaca ctccttctgg tgtgccgtac aacatgatta acattacttc gggcggcaac 540

gacggcgcca cgaccaacgg gctggccgtg accggtacct tggtgctgga gtggacgcgt 600

ctgtcggact tgactgggaa cgacgagtat gcccgcttga gccagagagc tgaagactac 660

cttctccacc cggagccagc gcagtacgaa ccgttccctg gattgattgg aagcgcagtc 720

aatattgccg acggcaagct cgccaatggt cacatcagct ggaatggtgg cgcagactcg 780

tactacgagt acctgatcaa gatgtacgtc tacgatcccg aacgctttgg cctctaccgg 840

gaccgctggg tcgcagctgc cgagtcgagc atcaaccatc tggcttcgca cccgtccacc 900

cgcccagacg tgactttctt ggccacttac aacgaggagc atcagctggg cctgaccagc 960

caacacctga cctgcttcga cggtggaagc tttctgcttg gtgggacatt gctggaccgc 1020

caggactttg tcgacttcgg ccttgacctt gtcgccggct gccacgagac gtacaactcg 1080

actctgacgg gcatcggccc tgagcaattc agctgggacc ctaacggtgt gcccgacagc 1140

cagaaggagc tgttcgagcg cgcaggcttc tacatcaaca gcggccaata cattcttcgt 1200

cccgaagtca tcgagagctt ctactatgca tggcgcgtca caggtgatgg aacgtacctc 1260

gaatgggtgt ggaacgcctt caccaacatc aacaagtact gccgcactgc gaccggtttc 1320

gcggggctgg agaacgtcaa tgcagcgaac ggcggaggcc ggatcgacaa ccaggagagt 1380

ttcatgttcg cagaggtgct gaagtattcg tttttgactt ttgctcctga ggacgactgg 1440

caggtgcaga agggcagtgg aaatacgttt gtttataaca ccgaggcgca cccgtttaag 1500

gtgtatacgc ctcag 1515

<210> 6
<211> 505
<212> PRT
<213> Aspergillus nidulans

<400> 6

Met Arg Thr Leu Leu Ala Leu Ala Ala Phe Ala Gly Phe Ala Ala Ala
1 5 10 15

Arg Val Pro Ala Thr Ala Ile Thr Arg Pro Val Met Arg Ser Asp Ser
20 25 30

Arg Ala Asp Ala Val Lys Glu Ala Phe Ser His Ala Trp Asp Gly Thr
35 40 45

Thr Asn Thr Ala Phe Pro His Asp Glu Leu His Pro Ile Ser Asn Gly
50 55 60

Thr Gly Asp Ser Arg Asn His Trp Gly Ala Ser Ala Val Asp Ala Leu
65 70 75 80

Ser Thr Ala Ile Met Met Arg Asn Ala Thr Ile Val Asn Gln Ile Leu
85 90 95

Asp His Ile Ala Ala Val Asp Thr Ser Lys Thr Asn Ala Met Val Ser
100 105 110

Leu Phe Glu Thr Thr Ile Arg Thr Leu Ala Gly Met Ile Ser Gly Thr
115 120 125

Asp Leu Leu Lys Gly Pro Ala Ala Gly Leu Val Asp Asp Ser Arg Val
130 135 140

```
Asp Val Leu Leu Glu Gln Ser Gln Asn Leu Ala Glu Val Leu Lys Phe
145                 150                 155                 160

Ala Phe Asp Thr Pro Ser Gly Val Pro Thr Asn Met Ile Asn Ile Thr
                165                 170                 175

Ser Gly Gly Asn Asp Gly Ala Thr Thr Asn Gly Leu Ala Val Thr Gly
            180                 185                 190

Thr Leu Val Leu Glu Trp Thr Arg Leu Ser Asp Leu Thr Gly Asn Asp
        195                 200                 205

Glu Thr Ala Arg Leu Ser Gln Arg Ala Glu Asp Thr Leu Leu His Pro
    210                 215                 220

Glu Pro Ala Gln Thr Glu Pro Phe Pro Gly Leu Ile Gly Ser Ala Val
225                 230                 235                 240

Asn Ile Ala Asp Gly Lys Leu Ala Asn Gly His Ile Ser Trp Asn Gly
                245                 250                 255

Gly Ala Asp Ser Thr Thr Glu Thr Leu Ile Lys Met Thr Val Thr Asp
            260                 265                 270

Pro Glu Arg Phe Gly Leu Thr Arg Asp Arg Trp Val Ala Ala Ala Glu
        275                 280                 285

Ser Ser Ile Asn His Leu Ala Ser His Pro Ser Thr Arg Pro Asp Val
    290                 295                 300

Thr Phe Leu Ala Thr Thr Asn Glu Glu His Gln Leu Gly Leu Thr Ser
305                 310                 315                 320

Gln His Leu Thr Cys Phe Asp Gly Gly Ser Phe Leu Leu Gly Gly Thr
                325                 330                 335

Leu Leu Asp Arg Gln Asp Phe Val Asp Phe Gly Leu Asp Leu Val Ala
            340                 345                 350

Gly Cys His Glu Thr Thr Asn Ser Thr Leu Thr Gly Ile Gly Pro Glu
        355                 360                 365
```

Gln Phe Ser Trp Asp Pro Asn Gly Val Pro Asp Ser Gln Lys Glu Leu
    370   375   380

Phe Glu Arg Ala Gly Phe Thr Ile Asn Ser Gly Gln Thr Ile Leu Arg
385   390   395   400

Pro Glu Val Ile Glu Ser Phe Thr Thr Ala Trp Arg Val Thr Gly Asp
                405   410   415

Gly Thr Thr Leu Glu Trp Val Trp Asn Ala Phe Thr Asn Ile Asn Lys
            420   425   430

Thr Cys Arg Thr Ala Thr Gly Phe Ala Gly Leu Glu Asn Val Asn Ala
        435   440   445

Ala Asn Gly Gly Gly Arg Ile Asp Asn Gln Glu Ser Phe Met Phe Ala
    450   455   460

Glu Val Leu Lys Thr Ser Phe Leu Thr Phe Ala Pro Glu Asp Asp Trp
465   470   475   480

Gln Val Gln Lys Gly Ser Gly Asn Thr Phe Val Thr Asn Thr Glu Ala
                485   490   495

His Pro Phe Lys Val Thr Thr Pro Gln
            500   505

<210> 7
<211> 6
<212> DNA
<213> Aspergillus nidulans

<400> 7
gtaagt 6

<210> 8
<211> 6
<212> DNA
<213> Aspergillus nidulans

<220>
<221> variation
<222> (4)..(4)
<223> N = A, C, G, or T

<400> 8
gtangt 6

<210> 9
<211> 6
<212> DNA
<213> Aspergillus nidulans

<400> 9
gctgac 6

<210> 10
<211> 6
<212> DNA
<213> Aspergillus nidulans

<400> 10
rctrac 6

<210> 11
<211> 6
<212> DNA
<213> Aspergillus nidulans

<400> 11
gtacgt 6

<210> 12
<211> 6
<212> DNA
<213> Aspergillus nidulans

<400> 12
actgac 6

<210> 13
<211> 10
<212> PRT
<213> Asperfillus nidulans

<400> 13

Gly Gly Leu Gly Glu Ser Phe Thr Glu Thr
1 5 10

<210> 14
<211> 9
<212> PRT
<213> Aspergillus nidulans

<400> 14

Leu Ala Glu Thr Leu Lys Thr Leu Thr
1 5

<210> 15
<211> 29
<212> DNA
<213> PCR Primer

<220>
<221> y represents c, t, or u; r represents g or a; n represents a, c, g, t, or u; and h represents a, c, t, or u
<222> (1)..(29)
<223>

<400> 15
ggyggyctng gygartcntt ctacgagta 29

<210> 16
<211> 33
<212> DNA
<213> PCR Primer

<220>
<221> y represents c, t, or u; r represents g or a; n represents a, c, g, t, or u; and h represents a, c, t, or u
<222> (1)..(33)
<223>

<400> 16
gtanaggtac ttnagngtct cngcnagrha gaa 33

<210> 17
<211> 2032
<212> DNA
<213> Aspergillus nidulans

<400> 17
atcgatgtct gctgcataaa ggcagacgga ggaagatgcc gagacggtgg tcctccctca 60

tcagcatcac agccatcttc ttggtcctct tcttcctcct tcataggaat acagacacac 120

cacgcgccgc caatagggct acaaacggcc ctgccaacgg ctttgctagg cagcaaagca 180

tatgtccatc aacaccccct cagcctccat ataaccgaac cagcacggga gggttcaact 240

ggggtgaaat cccagtcaga taccctgtat ccgacttcat cccgctgtca accaactctc 300

ctgcaacact tccgcgcatc caacgctctt ccttcccact tcaatcctca atcactaaat 360

cccgccaggc agcagtcaaa ggtgcctttc agcgcgcatg gacctcctac acaacccacg 420

cctggaaggc ggacgaggta cggcccatca cggccggatc tcgaaacaac tttggcggat 480

ggggagcgac cctagtcgac aatctcgaca cactgctaat catggggctg gacgaggagt 540

tcgcagcggc agtcgacgcg ctcgcagata tagaattcag cccgcactcg tccccatcct 600

cctcccagag cacaatcaac atattcgaaa cgacaatccg gtatctgggc ggcttgctcg 660

cggcgtatga tctcactggc tgtcgagaga ctcggctgct ggacaaagca atccagcttg 720

gggagatgat ctacacctcc ttcgacacag agaaccgcat gcccgtacca cggtggaatc 780

tgcacaaagc aggcaacgga gagcctcagc gcgcggcagt gcagggcgtg ctcgctgaac 840

tcgccagcag cagtctcgag ttcacgcggc tgtcgcagct gacgggggat atgcggtatt 900

tcgatgcggc atcccgcatt accgatctgc ttgactccca agccggccat acccggatcc 960

cggggttgtg gccagtcagc gtgaacctgc agaaaggcga tctgacccgt gggtcgacat 1020

tcagttttgg cgggatggcc gatagcgcct acgagtatct cggcaagacg tatcggctcc 1080

tcggtggtgt ggggaaaggg ccacagtacg agcgtctggc gcgaaacgca ctagatgccg 1140

ggattcgaca tctcctcttc cgaccgatga cgcctgatca tgcagatatc ctcctacccg 1200

gggtcgcgca cgcaaccagc tcttccgtgg gactcgagcc ccggacagag catctcgcct 1260

gttttgtggg tgggatgtac gcgctcgccg ggaagctttt ctcaaaccag acgtacctcg 1320

acaccggccg gaagctgaca gacggttgta tctggtacta cgataattca ccgctaggta 1380

tcatgccgga gatgttcacc gtgccggctt gtccgtcagt ggctgaatgt ccttgggacg 1440

aaacaagggg tggtatctac acctacgtgc gtgatgggca ctactttctg cgtcctgagg 1500

caatggagag tatcttctat atgtggcgca ttacagggga cgaaaagtac cgcgaggctg 1560

catggagaat gttcacggct atcgaagcgg ttacaaagac ggagtttggg aatgcggcgg 1620

tgcgggatgt tatggttgag gaaggaaatg taaagagaga agatagcatg gagagtttct 1680

ggatggcaga gacgttgaag tatctgtatc tgatatttgg ggagaccgat ttggtcagct 1740

tggacgactg ggtgttcaat acggaggcgc acccttttgag gggtgcaggg agttgacatt 1800

gtattcacac atcggtatag acaaattata gagtagacgt tcaaaacggc caaaactgaa 1860

tggatagact ccatatgcat tgaatataca atgtattcgc tgcaaagcat ggataaaata 1920

aagatgtaca aagtgtcttt gttgtcgctt tgaaagtggt atatcatccc atcataaggt 1980

ggcagtgtaa ccaaccctct atatcaccta catagacagc tgatagaccg gc 2032

<210> 18
<211> 586

<212> PRT
<213> Aspergillus nidulans

<400> 18

Met Pro Arg Arg Trp Ser Ser Leu Ile Ser Ile Thr Ala Ile Phe Leu
1 5 10 15

Val Leu Phe Phe Leu Leu His Arg Asn Thr Asp Thr Pro Arg Ala Ala
20 25 30

Asn Arg Ala Thr Asn Gly Pro Ala Asn Gly Phe Ala Arg Gln Gln Ser
35 40 45

Ile Cys Pro Ser Thr Pro Pro Gln Pro Pro Thr Asn Arg Thr Ser Thr
50 55 60

Gly Gly Phe Asn Trp Gly Glu Ile Pro Val Arg Thr Pro Val Ser Asp
65 70 75 80

Phe Ile Pro Leu Ser Thr Asn Ser Pro Ala Thr Leu Pro Arg Ile Gln
85 90 95

Arg Ser Ser Phe Pro Leu Gln Ser Ser Ile Thr Lys Ser Arg Gln Ala
100 105 110

Ala Val Lys Gly Ala Phe Gln Arg Ala Trp Thr Ser Thr Thr Thr His
115 120 125

Ala Trp Lys Ala Asp Glu Val Arg Pro Ile Thr Ala Gly Ser Arg Asn
130 135 140

Asn Phe Gly Gly Trp Gly Ala Thr Leu Val Asp Asn Leu Asp Thr Leu
145 150 155 160

Leu Ile Met Gly Leu Asp Glu Glu Phe Ala Ala Ala Val Asp Ala Leu
165 170 175

Ala Asp Ile Glu Phe Ser Pro His Ser Ser Pro Ser Ser Ser Gln Ser
180 185 190

Thr Ile Asn Ile Phe Glu Thr Thr Ile Arg Thr Leu Gly Gly Leu Leu
195 200 205

```
Ala Ala Thr Asp Leu Thr Gly Cys Arg Glu Thr Arg Leu Leu Asp Lys
    210                 215                 220

Ala Ile Gln Leu Gly Glu Met Ile Thr Thr Ser Phe Asp Thr Glu Asn
225                 230                 235                 240

Arg Met Pro Val Pro Arg Trp Asn Leu His Lys Ala Gly Asn Gly Glu
                245                 250                 255

Pro Gln Arg Ala Ala Val Gln Gly Val Leu Ala Glu Leu Ala Ser Ser
            260                 265                 270

Ser Leu Glu Phe Thr Arg Leu Ser Gln Leu Thr Gly Asp Met Arg Thr
        275                 280                 285

Phe Asp Ala Ala Ser Arg Ile Thr Asp Leu Leu Asp Ser Gln Ala Gly
    290                 295                 300

His Thr Arg Ile Pro Gly Leu Trp Pro Val Ser Val Asn Leu Gln Lys
305                 310                 315                 320

Gly Asp Leu Thr Arg Gly Ser Thr Phe Ser Phe Gly Gly Met Ala Asp
                325                 330                 335

Ser Ala Thr Glu Thr Leu Gly Lys Thr Thr Arg Leu Leu Gly Gly Val
            340                 345                 350

Gly Lys Gly Pro Gln Thr Glu Arg Leu Ala Arg Asn Ala Leu Asp Ala
        355                 360                 365

Gly Ile Arg His Leu Leu Phe Arg Pro Met Thr Pro Asp His Ala Asp
    370                 375                 380

Ile Leu Leu Pro Gly Val Ala His Ala Thr Ser Ser Ser Val Gly Leu
385                 390                 395                 400

Glu Pro Arg Thr Glu His Leu Ala Cys Phe Val Gly Gly Met Thr Ala
                405                 410                 415

Leu Ala Gly Lys Leu Phe Ser Asn Gln Thr Thr Leu Asp Thr Gly Arg
            420                 425                 430

Lys Leu Thr Asp Gly Cys Ile Trp Thr Thr Asp Asn Ser Pro Leu Gly
```

435 440 445

Ile Met Pro Glu Met Phe Thr Val Pro Ala Cys Pro Ser Val Ala Glu
450 455 460

Cys Pro Trp Asp Glu Thr Arg Gly Gly Ile Thr Thr Thr Val Arg Asp
465 470 475 480

Gly His Thr Phe Leu Arg Pro Glu Ala Met Glu Ser Ile Phe Thr Met
485 490 495

Trp Arg Ile Thr Gly Asp Glu Lys Thr Arg Glu Ala Ala Trp Arg Met
500 505 510

Phe Thr Ala Ile Glu Ala Val Thr Lys Thr Glu Phe Gly Asn Ala Ala
515 520 525

Val Arg Asp Val Met Val Glu Glu Gly Asn Val Lys Arg Glu Asp Ser
530 535 540

Met Glu Ser Phe Trp Met Ala Glu Thr Leu Lys Thr Leu Thr Leu Ile
545 550 555 560

Phe Gly Glu Thr Asp Leu Val Ser Leu Asp Asp Trp Val Phe Asn Thr
565 570 575

Glu Ala His Pro Leu Arg Gly Ala Gly Ser
580 585

<210> 19
<211> 6
<212> DNA
<213> Artificial Sequence

<220>
<223> Consensus Splice Site

<220>
<221> variation
<222> (1)..(1)
<223> R = G or A

<220>
<221> variation
<222> (4)..(4)
<223> R = G or A

<400> 19
rctrac 6

AGGGCCTTTTCAGGGGCTGGTTGGCCAAACCTGCCG
TATTTTCGATTCAGGGCGGGCCCTTAATTCTGGGGA
CCGATATCCCTTGAACGCGGGGGCAATCAGTTCAAC
CAACCCAGACCGCTGGGCTTGGTATGAGATTGCGAG
CATCTATGTCGGTGTACCTTCTGAATGACAATGAAT
GTATTTTACTTCTCGAAAAGAACCCTTGGGCACTGA
ATTGTGCGGAGAATGATGCCCTGATTATGATACAAC
TAGTCCGCTCCGTCAAGCCACAAGGGTCTGGGCAGT
CCGCTATAAATCAAAATCGCCTGCACGAACAGACGA
ATAACCAAGAAAACGCCCGAGCGCGAGCGTTTCTTC
TTCCTCTAAGCCTTGCAGCTGGCTCTGCGTCTTTGA
TCAACCCTTTAGCTGAATTTCCCCAGAACTTCAGCC
CTCTGCATCCTGTCCTTACCGCAACTCGTTAACCTG
CGCGACCTCGCGCGACCACAGCCTTAGGTTTCGAGA
TGCCATGAAAATCAGAAATTGAACCCCCTTTCCATT
ACTATCATTCTCTGCATTCTGCGAGTGATCTGTCCT
TCGACGTTCCTTCTTTCCAGCGCTGCGGCGCCTTCA
CTCTCGTTGCCTACGTTTGACCACGGTCCTACCTCT
CCTACTGCTGATTATTAGGCTCCTCCCTACGCCTCC
AATACAGGGAAGTCGCCGGCC

M F R A R R S R I S L V
F A V
ATGTTTCGTGCACGACGATCTCGCATCTCGCTGGTG
TTTGCCGTT

I F V L L I F H F S R L
A V T
ATATTTGTCCTCCTCATATTCCACTTTAGCCGTCTC
GCAGTTACG

Figure 1A

I S L Q S W V P P P P V
D H H
ATCAGCCTGCAATCTTGGGTACCTCCGCCGCCCGTC
GATCACCAT

N P P F P D Q N L K D P
Y E N
AATCCCCCTTTCCCCGACCAGAACCTCAAAGATCCA
TACGAAAAC

D N S A T G S G A P P P
A L V
GACAATAGTGCGACCGGCAGTGGGGCTCCTCCGCCT
GCGTTGGTA

E P E E Y Q R P P L Y T
D S D
GAGCCAGAAGAATACCAACGACCACCACTTTACACA
GATTCAGAT

D S P T P S K E R L D T
P S N
GACAGCCCAACTCCGTCAAAAGAACGCCTGGACACC
CCGAGCAAT

V P S Q E P E F D A A R
L Q T
GTCCCATCTCAGGAGCCTGAATTTGATGCCGCCAGA
CTTCAGACG

Figure 1B

```
 G  A  Q  T  Q  N  K  H  E  D  D  E
D  I  V
GGTGCGCAGACCCAAAATAAACATGAAGATGATGAG
GATATTGTC

 P  I  S  H  W  K  P  M  P  E  R  H
P  V  S
CCAATTTCTCACTGGAAGCCGATGCCCGAACGGCAT
CCAGTCAGT

 P  E  A  L  I  K  L  P  T  G  Q  S
K  E  L
CCGGAGGCTTTGATCAAGCTGCCAACCGGGCAATCA
AAGGAACTC

 P  Q  L  Q  A  K  F  K  D  E  S  S
S  D  K
CCCCAACTGCAAGCTAAGTTCAAGGACGAGTCGTCC
TCGGACAAG

 M  Q  R  L  Q  Q  L  D  T  I  K  S
A  F  L
ATGCAGCGGCTGCAACAACTTGACACTATCAAGTCG
GCGTTCTTA

 H  A  W  N  G  Y  K  I  S  A  M  G
H  D  E
CATGCGTGGAACGGTTACAAGATCTCTGCCATGGGT
CATGATGAG
```

Figure 1C

```
 V  R  P  L  R  G  G  F  K  D  T  F
N  G  W
GTTAGACCTCTGCGCGGTGGTTTCAAGGACACATTC
AATGGCTGG

 G  A  T  L  V  D  A  L  D  T  L  W
I  M  D
GGCGCGACCCTTGTCGACGCCTTGGATACCCTGTGG
ATCATGGAT

 L  K  E  E  F  S  M  A  V  D  Y  V
K  K  I
CTCAAAGAGGAGTTCTCCATGGCAGTCGACTACGTC
AAGAAAATC

 D  F  T  T  S  T  K  K  E  I  P  V
F  E  T
GATTTTACCACCAGCACCAAGAAAGAGATTCCGGTC
TTTGAAACC

 T  I  R  Y  L  G  G  M  L  G  A  Y
D  I  S
ACTATTCGCTACCTAGGCGGGATGCTCGGGGCCTAT
GATATTTCG

 G  H  K  Y  D  I  L  L  E  K  S  V
E  L  A
GGACACAAATACGATATACTTTTGGAAAAGTCTGTT
GAGCTTGCG
```

Figure 1D

```
 K  R  S  A  G  P  E  K  G  N  W  H
V  V  A
GATGTCTTGATGGACGCCTTCGACACACCGAACCGG
ATGCCAACC

 T  D  V  L  M  D  A  F  D  T  P  N
R  M  P
CTCTATTATAAATGGAGCCCAGAGTATGCTTCAGAG
TTTCGCCGG

 T  L  Y  Y  K  W  S  P  E  Y  A  S
E  F  R
GGGGACTTTAAGGCTGTTCTCGCCGAGCTTGGCTCT
CTCTCTCTC

 R  G  D  F  K  A  V  L  A  E  L  G
S  L  S
GAGTTCACGCGTTTGGCGCAGTTGACCAAACAGGAC
AAGTACTAC

 L  E  F  T  R  L  A  Q  L  T  K  Q
D  K  Y
GATGCAATTGCACGAATCACAAATGAGCTCGAAAAG
TATCAGGAT

 Y  D  A  I  A  R  I  T  N  E  L  E
K  Y  Q
TTGACAAAGCTTCCCGGCTTGTGGCCTCTCAACCTG
GACGCATCC
```

Figure 1E

```
 D   L   T   K   L   P   G   L   W   P   L   N
L   D   A
GGGTGCAGGCGAGTTCCCGGCGTCTCGCGAGAGCCT
GCTGCGGCT

 S   G   C   R   R   V   P   G   V   S   R   E
P   A   A
GGGCAGCCAGTCAGATGGTCCTCTGACGAGATCAAC
TCGACGAGC

 A   G   Q   P   V   R   W   S   S   D   E   I
N   S   T
TCGGTATCGTATCGTACAAGACAAATTCATGAGGGC
GGAGAGCCT

 S   S   V   S   Y   R   T   R   Q   I   H   E
G   G   E
GTCCGTCATGACAATGATTCGTTTGAAACGGGTTTT
CCTGTATCA

 P   V   R   H   D   N   D   S   F   E   T   G
F   P   V
GTCGATACTCGGACTCCTCCCCCAAAGCAAGATTGC
ACCGGAGGC

 S   V   D   T   R   T   P   P   P   K   Q   D
C   T   G
CTCAACGATCAGCTCTCAGGCATTGACAAGTTCGGA
CTCGGAGCC
```

Figure 1F

G L N D Q L S G I D K F G L G
CTTGGTGACTCTACGTACGAGTACTTACCGAAAGAG TATATGTTG

A L G D S T Y E Y L P K E Y M
CTCGGCGGTAACAACGACCAGTACCTCAACATGTAT CAGAAGGCC

L L G G N N D Q Y L N M Y Q K
ATGGACACAGTGCGAGAATATCTTGTTTATCAGCCA ATGCTCAAG

A M D T V R E Y L V Y Q P M L
AATAATCGCGATGTCCGCTTCTTAGCGACAGTTAGT ATGACAAAG

K N N R D V R F L A T V S M T
AGCCTTGATGCAAAACCTCCGGGGCGTACCACTTTC GCGTACGAA

K S L D A K P P G R T T F A Y
GGCACTCACCTCACCTGTTTTGCTGGTGGTATGCTT GCCATTGGC

Figure 1G

```
 E  G  T  H  L  T  C  F  A  G  G  M
L  A  I
GCCAAGTTGTTTGGGCTTGATAAGGATCTAAAGCTG
GGTAGTCAA

 G  A  K  L  F  G  L  D  K  D  L  K
L  G  S
CTGACGGACGGCTGTGTCTGGGCATATGAAGCCACA
AAGTCCGGA

 Q  L  T  D  G  C  V  W  A  Y  E  A
T  K  S
ATCATGCCGGAAGCATTCCAACTGGTCCCTTGTAAG
AAAGGCGAG

 G  I  M  P  E  A  F  Q  L  V  P  C
K  K  G
CCATGCGAATGGGATGAGGACGCATACTACATGGCC
ATGGATCCT

 E  P  C  E  W  D  E  D  A  Y  Y  M
A  M  D
TATGCCGACAAGCGGCCAATATCACATAACAAACGC
TCCGCCGGC

 P  Y  A  D  K  R  P  I  S  H  N  A
E  S  S
CCTGAAAAGGGGAATTGGCACGTCGTCGCCACAGCC
GAATCGTCT
```

Figure 1H

```
 S   P   Q   E   D   K   T   Q   K   S   T   T
T   E   G
TCGCCCCAGGAAGATAAAACACAGAAATCAACCACT
ACTGAGGGT

 R   H   T   G   T   T   T   G   A   G   A   L
S   H   E
CGACACACCGGTACAACTACCGGGGCAGGCGCGCTC
TCGCACGAG

 E   F   V   T   G   K   I   L   N   D   R   L
P   P   G
GAATTCGTCACGGGAAAAATCCTCAACGACCGACTC
CCGCCGGGC

 M   T   G   I   S   A   R   Q   Y   L   L   R
P   E   A
ATGACAGGGATCTCGGCTCGGCAGTACCTCCTTCGC
CCGGAGGCG

 I   E   S   V   F   I   M   F   R   L   T   G
D   P   S
ATCGAGTCTGTCTTCATCATGTTCCGCCTCACGGGC
GATCCTTCC

 W   R   E   K   G   W   K   M   F   Q   A   V
D   K   A
TGGCGCGAAAAGGGTTGGAAGATGTTCCAGGCTGTC
GACAAAGCC
```

Figure 1I

```
 T   K   T   E   L   A   N   S   A   I   S   D
V   T   V
ACGAAGACGGAGCTGGCGAACTCGGCCATTTCCGAC
GTAACCGTC

 D   N   P   R   P   V   D   S   M   E   S   F
W   L   A
GATAATCCACGCCCGGTGGACAGTATGGAATCATTC
TGGCTTGCG

 E   T   L   K   Y   F   Y   L   L   F   S   D
P   S   L
GAGACTCTGAAATACTTCTACCTTCTTTTCAGCGAT
CCAAGCCTG

 V   S   L   D   E   Y   V   L
GTGAGCCTTGAGGAATATGTCTTGTAAGTGATGCTT
GACTTAATC

 intron 1                                N   T
E   A   H
GACTGCTTGATGCTGACTTTTCCCTTAGGAACACCG
AGGCTCATC

P   F   K   R   P   K   Y  ***
CGTTCAAGCGACCCAGGTACTGAAGTACTAATTTAA
ATGATCTTTTAGCCTGTATCTATACATGGCCGCTCC
GCTGTAGAAGCATTGATACCATTAAGACAGTATCGC
TGCATTCGTGTACCATTTGAGCTTCCAGAGGAACCT
CTTT
```

Figure 1J

TACCATGATCGACAGGTTGTTGCCCATATTGGG
ACTAGATTTGTTTACCCCAGTCGGAATGTGCCT
AAAGTGGAAGGTATGATGATGCCCAGGATCGCG
CCCCAGTCATCAACTCCATCATGGGACGGTCCT
TGATCCTCAAGGCACGAAGTGGAGATCAGGTCC
GTAGTGCATATGCATGGCCCATCAGCCTGAAGC
ACTTCCCCAAGCAAAGTCGAGACTCGGACACCG
ATGATATCCCTGCTGTCCCTGACTGATGCATAG
TGCATGCCCCTGCGCTGGCTCCCCTTTTCACTC
CGCCTGGTCTCCAGTCTCCACTCCTCACCATTG
ATGTCTGCCCCCGCCCGCCCTCCATCCTCCATC
ATTCTTATATCTACGGACTCGGTCACTCGTTAT
CACTAGAGTCCTTGTTTATTCCTAGTGTTTGCA
TTCTTACGTGTAGTT

M R T L L A L A A F A
G F A A
ATGCGTACGCTTCTCGCTCTCGCGGCCTTCGCG
GGCTTTGCCGCT

A R V P A Y A I T R P
V M R S
GCTAGGGTGCCCGCCTACGCCATCACGCGCCCG
GTGATGCGCAGT

D S R A D A V K E A F
S H A W
GATTCTCGCGCCGACGCTGTCAAGGAGGCCTTT
TCGCATGCCTGG

Figure 2A

```
 D  G  Y  Y  N  Y  A  F  P  H  D
E  L  H  P
GACGGTTACTACAACTACGCTTTTCCTCATGAC
GAGCTTCACCCG

 I  S  N  G  Y  G  D  S  R  N  H
W  G  A  S
ATTTCGAACGGTTACGGAGACTCGCGAAACCAC
TGGGGCGCGTCG

 A  V  D  A  L  S  T  A  I  M  M
R  N  A  T
GCCGTCGACGCTCTATCGACGGCCATCATGATG
CGCAACGCGACC

 I  V  N  Q  I  L  D  H  I  A  A
V  D  Y  S
ATCGTCAACCAGATCCTTGACCATATTGCTGCT
GTGGACTACTCC

 K  T  N  A  M  V  S  L  F  E  T
T  I  R  Y
AAGACCAACGCCATGGTAAGTTTGTTCGAGACG
ACCATCCGGTAC

 L  A  G  M  I  S  G  Y  D  L  L
K  G  P  A
CTCGCGGGCATGATTTCTGGATACGACCTGCTC
AAAGGCCCTGCG
```

Figure 2B

```
 A   G   L   V   D   D   S   R   V   D   V
L   L   E   Q
GCGGGGTTGGTGGACGACAGCAGGGTCGACGTG
CTTCTAGAGCAG

 S   Q   N   L   A   E   V   L   K   F   A
F   D   T   P
TCGCAGAACCTCGCCGAGGTGCTGAAATTCGCG
TTCGACACTCCT

 S   G   V   P   Y   N   M   I   N   I   T
S   G   G   N
TCTGGTGTGCCGTACAACATGATTAACATTACT
TCGGGCGGCAAC

 D   G   A   T   T   N   G   L   A   V   T
G   T   L   V
GACGGCGCCACGACCAACGGGCTGGCCGTGACC
GGTACCTTGGTG

 L   E   W   T   R   L   S   D   L   T   G
N   D   E   Y
CTGGAGTGGACGCGTCTGTCGGACTTGACTGGG
AACGACGAGTAT

 A   R   L   S   Q   R   A   E   D   Y   L
L   H   P   E
GCCCGCTTGAGCCAGAGAGCTGAAGACTACCTT
CTCCACCCGGAG
```

Figure 2C

```
 P   A   Q   Y   E   P   F   P   G   L   I
G   S   A   V
CCAGCGCAGTACGAACCGTTCCCTGGATTGATT
GGAAGCGCAGTC

 N   I   A   D   G   K   L   A   N   G   H
I   S   W   N
AATATTGCCGACGGCAAGCTCGCCAATGGTCAC
ATCAGCTGGAAT

 G   G   A   D   S   Y   Y   E   Y   L   I
K   M   Y   V
GGTGGCGCAGACTCGTACTACGAGTACCTGATC
AAGATGTACGTC

 Y   D   P   E   R   F   G   L   Y   R   D
R   W   V   A
TACGATCCCGAACGCTTTGGCCTCTACCGGGAC
CGCTGGGTCGCA

 A   A   E   S   S   I   N   H   L   A   S
H   P   S   T
GCTGCCGAGTCGAGCATCAACCATCTGGCTTCG
CACCCGTCCACC

 R   P   D   V   T   F   L   A   T   Y   N
E   E   H   Q
CGCCCAGACGTGACTTTCTTGGCCACTTACAAC
GAGGAGCATCAG
```

Figure 2D

```
 L   G   L   T   S   Q   H   L   T   C   F
D   G   G   S
CTGGGCCTGACCAGCCAACACCTGACCTGCTTC
GACGGTGGAAGC

 F   L   L   G   G   T   L   L   D   R   Q
D   F   V   D
TTTCTGCTTGGTGGGACATTGCTGGACCGCCAG
GACTTTGTCGAC

 F   G   L   D   L   V   A   G   C   H   E
T   Y   N   S
TTCGGCCTTGACCTTGTCGCCGGCTGCCACGAG
ACGTACAACTCG

 T   L   T   G   I   G   P   E   Q   F   S
W   D   P   N
ACTCTGACGGGCATCGGCCCTGAGCAATTCAGC
TGGGACCCTAAC

 G   V   P   D   S   Q   K   E   L   F   E
R   A   G   F
GGTGTGCCCGACAGCCAGAAGGAGCTGTTCGAG
CGCGCAGGCTTC

 Y   I   N   S   G   Q   Y   I   L   R   P
E   V   I   E
TACATCAACAGCGGCCAATACATTCTTCGTCCC
GAAGTCATCGAG
```

Figure 2E

S F Y Y A W R V T G D
G T
AGCTTCTACTATGCATGGCGCGTCACAGGTGAT
GGAACG**GTACGT**

intron 1
TCACTCAGCGCTGCTTCCGTAGGAAGACCATAC
TGACCGCTTTAG

Y L E W V W N A F T N
I N K Y
TACCTCGAATGGGTGTGGAACGCCTTCACCAAC
ATCAACAAGTAC

C R T A T G F A G L E
N V N A
TGCCGCACTGCGACCGGTTTCGCGGGGCTGGAG
AACGTCAATGCA

A N G G G R I D N Q E
S F M F
GCGAACGGCGGAGGCCGGATCGACAACCAGGAG
AGTTTCATGTTC

A E V L K Y S F L T F
A P
GCAGAGGTGCTGAAGTATTCGTTTTTGACTTTT
GCTCCTG**GTACG**

Figure 2F

intron 2

TTTCCCTAGTTCTGGTTCACCTGTGGAGAATAT
TACTGACTGCAG

E D D W Q V Q K G S
G N T F
CAGAGGACGACTGGCAGGTGCAGAAGGGCAGTG
GAAATACGTTTG

V Y N T E A H P F K V
Y T P Q *
TTTATAACACCGAGGCGCACCCGTTTAAGGTGT
ATACGCCTCAGT

**
AGATAGTACATATTGTGCTCTAGCGTACTGCAT
GCATAACCTAAATGGCATCACCTACTTACTGAC
TACTCTACTGACAAGCAGTTGCCTTTATATGCT
CTTAAGGAGCTTGGGCGGTAAATCGTAATTCAC
TCTAAACTCCGCGTGCTGTTCGGGGAAGGCTGG
GC

Figure 2G

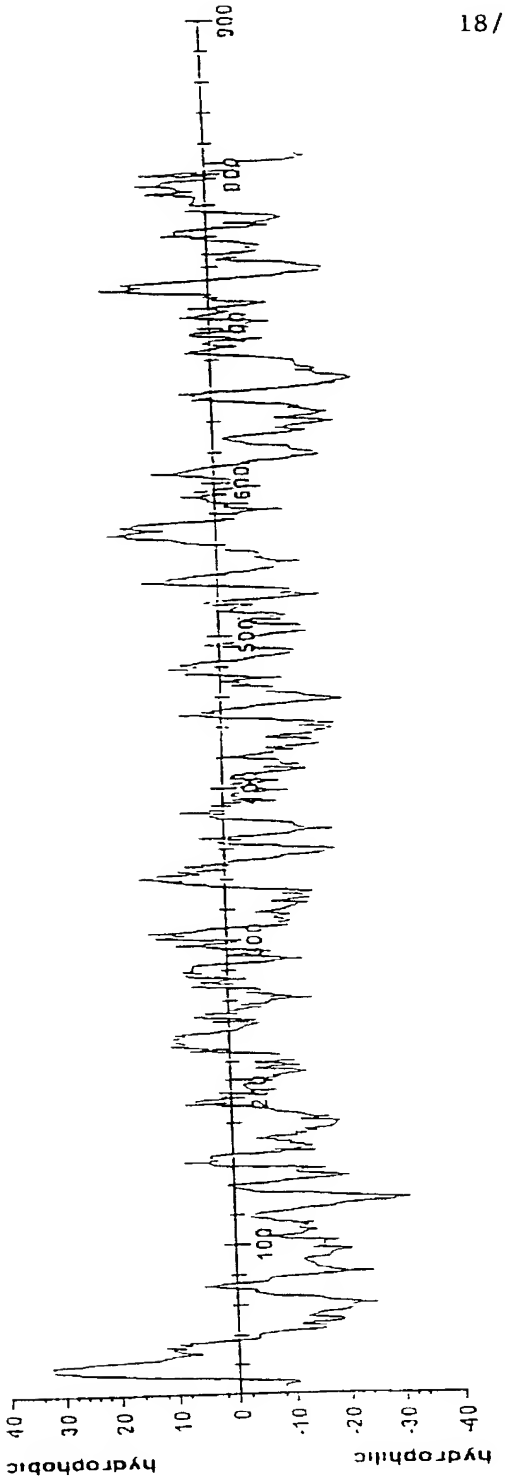

FIGURE 3A

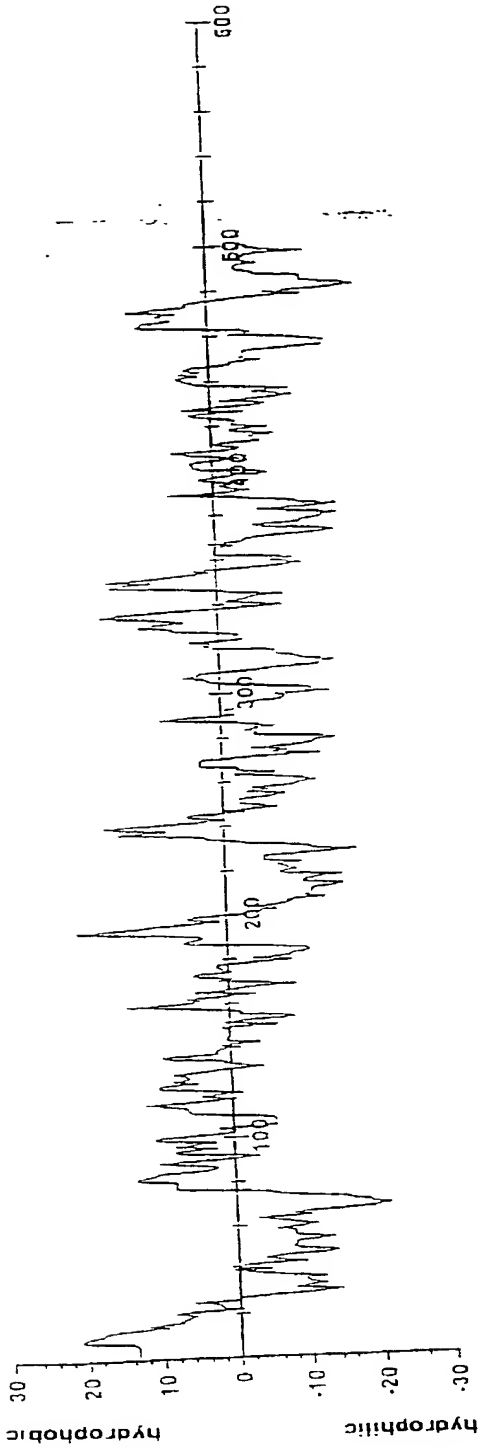

FIGURE 3B

	1	2	3	4	5	6	7	8	9	10	11	12	13
1-S.cerevisiae		24	24	22	22	21	27	27	28	28	29	28	27
2-A.nidulans 1A			38	28	30	31	23	26	25	25	25	25	24
3-O.novo-ulmi 1A				26	27	26	23	25	22	21	23	21	24
4-A.nidulans 1B					70	66	27	27	26	28	28	28	28
5-A.satoi						72	26	26	29	29	30	30	29
6-P.citrinum							25	26	26	28	28	28	27
7-D.melanogaster								49	42	43	42	43	43
8-S.frugiperda									42	41	43	41	42
9-Human 1A										58	85	57	82
10-Human 1B											60	94	60
11-Mouse 1A												60	85
12-Mouse 1B													58
13-S.scrofa													

FIGURE 4

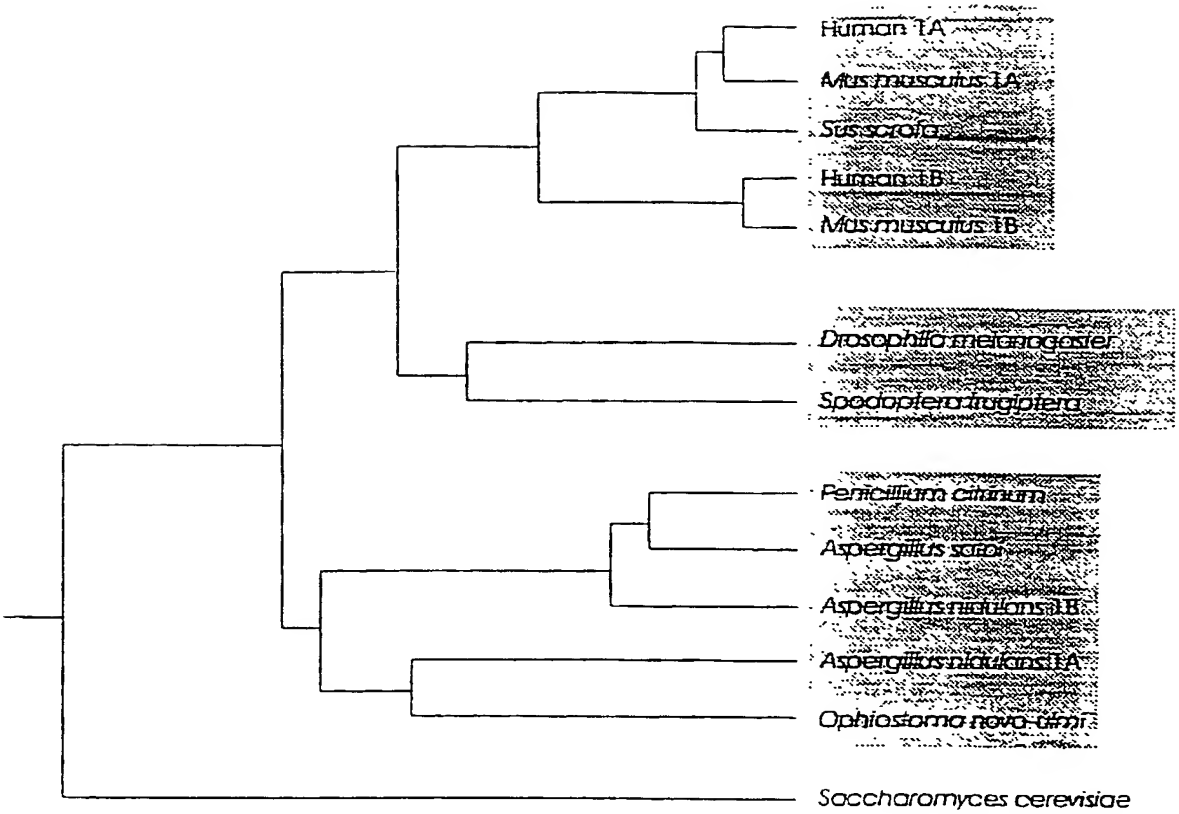
Human 1A
Mus musculus 1A
Sus scrofa
Human 1B
Mus musculus 1B
Drosophila melanogaster
Spodoptera frugiperda
Penicillium citrium
Aspergillus sato
Aspergillus nidulans 1B
Aspergillus nidulans 1A
Ophiostoma novo-ulmi
Saccharomyces cerevisiae

FIGURE 5

```
FEATURES             Location/Qualifiers
     source          1..2032
                     /organism="Emericella nidulans"
                     /db_xref="taxon:5072"
     gene            36..1796
                     /gene="MNSIC"
     CDS             36..1796
                     /gene="MNSIC"
                     /note=" [intronless gene]"
                     /codon_start=1
                     /product="alpha-mannosidase IC"

/translation="MPRRWSSLISITAIFLVLFFLLHRNTDTPRAANRATNGPANGF
ARQQSICPSTPPQPPYNRTSTGGFNWGEIPVRYPVSDFIPLSTNSPATLPRIQRSSFPLQSSITKS
RQAAVKGAFQRAWTSYTTHAWKADEVRPITAGSRNNFGGWGATLVDNLDTLLIMGLD
EEFAAAVDALADIEFSPHSSPSSSQSTINIFETTIRYLGGLLAAYDLTGCRETRLLD
KAIQLGEMIYTSFDTENRMPVPRWNLHKAGNGEPQRAAVQGVLAELASSSLEFTRLS
QLTGDMRYFDAASRITDLLDSQAGHTRIPGLWPVSVNLQKGDLTRGSTFSFGGMADS
AYEYLGKTYRLLGGVGKGPQYERLARNALDAGIRHLLFRPMTPDHADILLPGVAHAT
SSSVGLEPRTEHLACFVGGMYALAGKLFSNQTYLDTGRKLTDGCIWYYDNSPLGIMP
EMFTVPACPSVAECPWDETRGGIYTYVRDGHYFLRPEAMESIFYMWRITGDEKYREA
AWRMFTAIEAVTKTEFGNAAVRDVMVEEGNVKREDSMESFWMAETLKYLYLIFGETD
LVSLDDWVFNTEAHPLRGAGS"
BASE COUNT       491 a    554 c    555 g    432 t
ORIGIN
1 atcgatgtct gctgcataaa ggcagacgga ggaagatgcc gagacggtgg
tcctccctca
61 tcagcatcac agccatcttc ttggtcctct tcttcctcct tcataggaat
acagacacac
121 cacgcgccgc caatagggct acaaacggcc ctgccaacgg
ctttgctagg cagcaaagca
181 tatgtccatc aacacccccct cagcctccat ataaccgaac
cagcacggga gggttcaact
241 ggggtgaaat cccagtcaga taccctgtat ccgacttcat
cccgctgtca accaactctc
301 ctgcaacact tccgcgcatc caacgctctt ccttcccact
tcaatcctca atcactaaat
361 cccgccaggc agcagtcaaa ggtgcctttc agcgcgcatg
gacctcctac acaacccacg
421 cctggaaggc ggacgaggta cggcccatca cggccggatc
tcgaaacaac tttggcggat
481 ggggagcgac cctagtcgac aatctcgaca cactgctaat
catggggctg gacgaggagt
541 tcgcagcggc agtcgacgcg ctcgcagata tagaattcag
cccgcactcg tccccatcct
601 cctcccagag cacaatcaac atattcgaaa cgacaatccg
gtatctgggc ggcttgctcg
661 cggcgtatga tctcactggc tgtcgagaga ctcggctgct
ggacaaagca atccagcttg
```

Figure 6A

721 gggagatgat ctacacctcc ttcgacacag agaaccgcat
gcccgtacca cggtggaatc
781 tgcacaaagc aggcaacgga gagcctcagc gcgcggcagt
gcagggcgtg ctcgctgaac
841 tcgccagcag cagtctcgag ttcacgcggc tgtcgcagct
gacgggggat atgcggtatt
901 tcgatgcggc atcccgcatt accgatctgc ttgactccca
agccggccat acccggatcc
961 cggggttgtg gccagtcagc gtgaacctgc agaaaggcga
tctgacccgt gggtcgacat
1021 tcagttttgg cgggatggcc gatagcgcct acgagtatct
cggcaagacg tatcggctcc
1081 tcggtggtgt ggggaaaggg ccacagtacg agcgtctggc
gcgaaacgca ctagatgccg
1141 ggattcgaca tctcctcttc cgaccgatga cgcctgatca
tgcagatatc ctcctacccg
1201 gggtcgcgca cgcaaccagc tcttccgtgg gactcgagcc
ccggacagag catctcgcct
1261 gttttgtggg tgggatgtac gcgctcgccg ggaagctttt
ctcaaaccag acgtacctcg
1321 acaccggccg gaagctgaca gacggttgta tctggtacta
cgataattca ccgctaggta
1381 tcatgccgga gatgttcacc gtgccggctt gtccgtcagt
ggctgaatgt ccttgggacg
1441 aaacaagggg tggtatctac acctacgtgc gtgatgggca
ctactttctg cgtcctgagg
1501 caatggagag tatcttctat atgtggcgca ttacagggga
cgaaaagtac cgcgaggctg
1561 catggagaat gttcacggct atcgaagcgg ttacaaagac
ggagtttggg aatgcggcgg
1621 tgcgggatgt tatggttgag gaaggaaatg taaagagaga
agatagcatg gagagtttct
1681 ggatggcaga gacgttgaag tatctgtatc tgatatttgg
ggagaccgat ttggtcagct
1741 tggacgactg ggtgttcaat acggaggcgc acccttgag
gggtgcaggg agttgacatt
1801 gtattcacac atcggtatag acaaattata gagtagacgt
tcaaaacggc caaaactgaa
1861 tggatagact ccatatgcat tgaatataca atgtattcgc
tgcaaagcat ggataaaata
1921 aagatgtaca aagtgtcttt gttgtcgctt tgaaagtggt
atatcatccc atcataaggt
1981 ggcagtgtaa ccaaccctct atatcaccta catagacagc
tgatagaccg gc

Figure 6B

COMBINED DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled MANNOSIDASES AND METHODS FOR USING SAME, the specification of which

☐ is attached hereto.

☐ was filed on _____ as United States Patent Application No. _____.

☒ was described and claimed in PCT International Application No. PCT/US00/27210, filed on 2 October 2000, and as amended under PCT Articles 19 on n/a (if applicable).

☐ and was amended on _____ (if applicable).

☐ with amendments through _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 C.F.R § 1.56. If this is a continuation-in-part application filed under the conditions specified in 35 U.S.C. § 120 which discloses claims and subject matter in addition to that disclosed in the prior copending application, I further acknowledge the duty to disclose material information as defined in 37 C.F.R. § 1.56 which occurred between the filing date of the prior application and the national or PCT international filing date of the continuation-in-part application.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the applications(s) on which priority is claimed:

Number	Country	Day/Month/Year Filed	☐ Yes	☐ No

I hereby claim the benefit under 35 U.S.C. § 119(e) of any United States provisional application(s) listed below:

Application Number	Filing Date
60/157,341	October 1, 1999

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s) or § 365(c) of any PCT international application(s) designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, § 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

Application Number	Filing Date	Status: patented, pending abandoned

I hereby appoint the practitioners associated with the customer number provided below to prosecute this application, to file a corresponding international application, and to transact all business in the Patent and Trademark Office connected therewith:

Customer Number 

Name	Reg. No.	Name	Reg. No.
BUNKER, Gillian	47,461	ORR, David E.	44,988
BURG, Daniel B.	41,649	PETERSEN, David P.	28,106
CALDWELL, Lisa M.	41,653	POLLEY, Richard J.	28,107
CARLSON, Anne	47,472	RINEHART, Kyle B.	47,027
GIRARD, Michael P.	38,467	RUPERT, Wayne W.	34,420
HAENDLER, Jeffrey B.	43,652	RYBAK, Sheree L.	47,913
HARDING, Tanya M.	42,630	SCOTTI, Robert F.	39,830
JAKUBEK, Joseph T.	34,190	SIEGEL, Susan Alpert	43,121
JONCUS, Stephen J.	44,809	SLATER, Stacey C.	36,011
JONES, Michael D.	41,879	STEPHENS Jr., Donald L.	34,022
KLARQUIST, Kenneth S.	16,445	STUART, John W.	24,540
KLITZKE II, Ramon A.	30,188	VANDENBERG, John D.	31,312
LEIGH, James S.	20,434	WHINSTON, Arthur L.	19,155
MC LEOD, Richard D.	46,921	WIGHT, Stephen A.	37,759
MAURER, Gregory L.	43,781	WINN, Garth A.	33,220
MIRHO, Charles A.	41,199	ZASTROW, Devon J.	50,206
NOONAN, William D.	30,878		

I hereby grant the law firm of Klarquist Sparkman, LLP, the power to insert on this Combined Declaration and Power of Attorney any further identification which may be necessary or desirable in order to comply with the rules of the United States Patent and Trademark Office for submitting this document.

Address all telephone calls to Sheree Lynn Rybak, Ph.D. at telephone number (503) 226-7391.

Address all correspondence to the address associated with the customer number provided below:

Customer Number 

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1-00

Name of First or Sole Inventor: William E. Hintz
Residence: Victoria, BC, Canada CAX
Mailing Address: 5601 Matterhorn Court
Victoria, BC V9E 2G1
CANADA
Citizenship: Canada
Inventor's Signature William E. Hintz **Date** March 15, 2002

2-00

Name of Second Inventor: Caleb Joshua Eades
Residence: Victoria, BC, Canada CAX
Mailing Address: 852 Leslie Drive
Victoria, BC V8X 2Y4
CANADA
Citizenship: United States of America
Inventor's Signature Caleb Joshua Eades **Date** MARCH 15, 2002